



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Informatique et Télécommunication

Présentée et soutenue par :

M. DAMIEN MARIYANAYAGAM

le vendredi 12 juin 2020

Titre :

Localisation et reconnaissance de marqueurs circulaires à partir d'une vue
de leur contour

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de Recherche en Informatique de Toulouse (IRIT)

Directeur(s) de Thèse :

M. VINCENT CHARVILLAT

MME SYLVIE CHAMBON

Rapporteurs :

M. ADRIEN BARTOLI, UNIVERSITE CLERMONT-FERRAND 1

M. DAVID FOFI, UNIVERSITE DE BOURGOGNE

Membre(s) du jury :

Mme LUCE MORIN, INSA DE RENNES, Président

Mme SYLVIE CHAMBON, TOULOUSE INP, Membre

Mme YIHONG WU, CHINESE ACADEMY OF SCIENCES, Membre

M. PIERRE GURDJOS, TOULOUSE INP, Membre

M. VINCENT CHARVILLAT, TOULOUSE INP, Membre

Table des matières

Introduction	3
1 Introduction aux marqueurs fiduciaux	7
1.1 Définition du marqueur fiduciaire	8
1.2 Différents types de marqueurs plans	10
1.3 Détection d'un marqueur dans l'image	12
1.3.1 Prétraitement de l'image	12
1.3.2 Détection du contour du marqueur	13
1.3.3 Validation des candidats détectés	16
1.4 Estimation géométrique de la pose d'un marqueur	17
1.4.1 Estimation précise des paramètres du contour	17
1.4.2 Estimation de l'homographie	17
1.4.3 Calcul de la pose dans le cadre des marqueurs fiduciaux	18
1.5 Reconnaissance et identification d'un marqueur	18
1.6 Conclusion et présentation d'un marqueur spécifique, le bleam	19
2 Calcul de l'homographie pour les marqueurs circulaires	21
2.1 Calcul de l'homographie de rectification par sa décomposition	22
2.2 Méthodes d'estimation de l'homographie	23
2.2.1 À partir de points et de droites	24
2.2.2 À partir de plusieurs cercles ou ellipses coplanaires	25
2.2.3 À partir de cercles concentriques	30
2.2.4 À partir d'un cercle, dans le cas calibré	35
2.3 Application à un cas spécifique : le contour d'un bleam	36
2.3.1 Présentation des méthodes comparées	37
2.3.2 Résultats des tests sur la rectification	41
2.4 Conclusion sur les méthodes et perspectives	42
3 Rectification métrique d'un bleam vu par une caméra calibrée par défaut	45
3.1 Estimation statistique de l'image du centre	46
3.2 Méthode différentiable proposée	47
3.2.1 Fonctions de calcul de l'image du centre	47
3.2.2 Calcul de la normale du plan de support	48
3.2.3 Calcul des valeurs propres de la matrice de l'image du cercle	49
3.3 Propagation de l'erreur et influence de la focale	50
3.3.1 Propagation d'incertitude sur l'image du centre	50
3.3.2 Incertitude sur les paramètres d'entrée	51
3.3.3 Calcul de la fonction de répartition	51

TABLE DES MATIÈRES

3.4	Un critère de validation photométrique des candidats	52
3.4.1	Calcul de la similitude d'image	53
3.4.2	Similitude d'image par la mesure statistique des intensités	54
3.4.3	Recherche d'un critère invariant en rotation	56
3.4.4	Création de critères photométriques pour notre méthode	57
3.5	Une évaluation de notre méthode	59
3.5.1	Construction d'un calibrage par défaut	59
3.5.2	Évaluation des candidats générés	61
3.6	Conclusion et perspectives sur la méthode proposée	62
4	Étude de l'ambiguïté de la pose d'un cercle vu par une caméra	65
4.1	Ambiguïté de la pose du point de vue d'une caméra monoculaire	66
4.2	Nouvelle formulation du calcul de la pose d'un cercle	68
4.2.1	Calcul de la ligne de fuite à partir des coniques dégénérées	69
4.2.2	Une observation intéressante sur les points de base	73
4.3	Paramétrage minimal pour étudier l'ambiguïté	74
4.3.1	Description du paramétrage	74
4.3.2	Une première condition nécessaire et suffisante	75
4.3.3	Une condition suffisante exploitable	76
4.4	Conclusion sur l'étude de l'ambiguïté	77
5	Localisation d'un marqueur enrichi de son environnement	79
5.1	Combiner marqueurs et points naturels pour faciliter le calcul de pose et la reconnaissance	80
5.1.1	Motivation	80
5.1.2	État de l'art	82
5.2	Estimation de la pose du marqueur avec les correspondances de points .	84
5.2.1	Formulation du problème	84
5.2.2	Résolution du problème aux moindres carrés	88
5.3	Distribution des angles estimés à partir des correspondances	91
5.3.1	Résolution directe à partir d'une correspondance	91
5.3.2	Identification de la référence à partir de la répartition des angles	93
5.4	Résolution du problème avec RANSAC (RANdom SAmple Consensus) .	95
5.4.1	Méthode d'estimation robuste	95
5.4.2	Évaluation par distance à la droite épipolaire	96
5.5	Résultats	97
5.5.1	Description des tests	97
5.5.2	Résultats sur la reconnaissance des marqueurs	98
5.5.3	Résultats sur l'estimation de la pose	99
5.6	Conclusion et intérêt de la méthode proposée	100
	Conclusion	103
	Annexes	
	Annexe A Rappel de géométrie projective	111
A.1	Notation	112
A.2	Modèle perspectif de la caméra	112
A.3	Espace projectif et projection de primitive usuelle	113
A.4	Homographie et projection de plan	114

A.5	Coniques et leurs projections	114
A.6	Quelques propriétés des coniques	115
A.7	Transformations projectives et invariants	116
A.7.1	Invariant projectif	116
A.7.2	Décomposition de l'homographie	117
A.7.3	Invariants des transformations projectives	117
A.8	Rectification affine et rectification métrique	118
A.9	Application au cas d'une projection par une caméra	118
Annexe B Ajustement d'une ellipse à un ensemble de points		121
B.1	Formulation du problème	122
B.2	Résolution avec des distances alternatives	123
B.3	Ajustement d'une ellipse au maximum de vraisemblance	124
Annexe C Résultats expérimentaux supplémentaires		127
C.1	Description du protocole d'évaluation	128
C.2	Analyse de nos résultats	130
Annexe D Preuve en calcul formel <i>Matlab</i>		133
D.1	Preuve de la proposition 4.2.6	133
D.2	Preuve de la proposition 4.3.1	134

TABLE DES MATIÈRES

Abstract

Locating or estimating the pose of a camera from a view is an essential problem in many applications, such as tracking a trajectory of autonomous vehicle navigation or augmented reality. The pose of a camera can be estimated from markers built specifically for this task and which can be placed in the scene. These markers have characteristics that distinguish them from other objects in the scene, making them easier to detect and recognize than natural objects such as points of interest. The correspondence between the shape of the marker and its image by a camera gives constraints to the pose so that it is possible to find the coordinates of the marker in the camera's reference frame.

Circular markers are particularly interesting for their robustness to partial occlusion of their image. Indeed an ellipse can be estimated from only 5 points of its contour so a partially occluded marker can always be correctly located. The problem is that the image of the circular contour is not sufficient to calculate the marker pose. Consequently, the circular markers work either with a calibrated camera or by adding other circular contours such as concentric circles.

Unfortunately, in some applications, the contour image is the only reliable information available. the camera is calibrated, the image of a single circle is not sufficient to obtain the pose of its support plane. Indeed, there is a double ambiguity that cannot be solved without additional information. In addition to this ambiguity, there is an infinity of possible poses for the marker. These poses correspond to all the rotations about the axis perpendicular to the support plane and passing through the center of the circle. Thus, we can distinguish four major issues addressed in this thesis :

- How to take into account approximate knowledge of camera calibration parameters ?
- What are the minimum conditions for solving the ambiguity on the pose of the plane ?
- How to calculate the rotation around the axis of the circle ?
- Is it possible to identify the marker using the image of its contour and its surrounding ?

The first contribution that handles this first question, introduce what we name : a "default" model of the camera. This model let us incorporate the uncertainties about the intrinsic parameters of the camera into an algorithm that searches for a photometric solution to the problem of metric correction of the marker image. The results show that the use of this "default" model offers encouraging prospects for circular markers.

The second contribution concerns the ambiguity of the pose of the support plan. To study it, we propose a new formulation of the problem of estimating the pose of a circle from the image of its contour. This formulation allows to express the link between geometric ambiguity and the algebraic solutions obtained. This leads us to propose a minimal parameterization of the pose that provides a simple condition on the resolution of ambiguity.

Finally, we propose a new method for locating the circular markers based on points of interest detected in its vicinity. The method needs a library of marker reference views made offline. The marker is recognized in a new view by both using the image of the circle and the point correspondences found in the reference views. The pose is then validated by a RANSAC procedure based on a minimal parameterization by the image of a 3D point and a circle.

TABLE DES MATIÈRES

Introduction

Depuis les premiers dessins dans les grottes jusqu'à nos lunettes de réalité augmentée contemporaines, l'homme a toujours été fasciné par sa propre perception visuelle de l'environnement. Le but étant de tenter de la comprendre, mais surtout de tenter de la reproduire. La vision par ordinateur est au cœur de cette démarche et profite des avancées scientifiques actuelles à la fois en informatique et en mathématique pour répondre à ce besoin. Les premiers modèles perspectifs datent de l'antiquité et se sont surtout développés au cours de la renaissance avec les codes de la perspective cavalière. Il faudra cependant attendre la formalisation de la géométrie projective pour élaborer un modèle plus complet qui fixera les prémices de la vision par ordinateur telle que nous la connaissons aujourd'hui. Le développement des capacités informatiques et du traitement de l'image à la fin des années 90 a permis une grande avancée avec l'apparition des descripteurs locaux et le perfectionnement d'algorithmes de détection des formes. Il est devenu possible d'apparier automatiquement des points, des droites et des ellipses et de résoudre l'un des problèmes les plus importants de la vision par ordinateur : l'estimation de la pose d'un observateur à partir d'une vue.

La pose d'une caméra correspond aux paramètres de position et d'orientation de cette caméra dans l'espace. La localisation d'une caméra ou l'estimation de la pose consiste donc à retrouver les coordonnées de la position et de l'orientation d'une caméra par rapport à un repère attaché à la scène filmée. Ce problème est essentiel dans de nombreuses applications, telles que, le suivi d'une trajectoire, dans le cadre de la navigation d'un véhicule autonome, ou la superposition d'une scène virtuelle à une scène réelle filmée par une caméra, dans le cadre de la réalité augmentée. Dans ce dernier exemple, il est nécessaire de pouvoir estimer et suivre la pose de la caméra réelle afin que les deux contenus (réel et virtuel) restent cohérents l'un par rapport à l'autre. L'estimation de la pose de la caméra est un problème central pour ces deux types d'applications.

Ainsi, l'estimation de la pose de la caméra est un sujet très étudié et il existe de nombreuses méthodes pour résoudre ce problème. De plus, chaque cadre applicatif implique ses propres contraintes qui diversifient le problème concret à résoudre. À titre d'exemple, si les données de capteurs inertiels ou GPS assistent la localisation visuelle ou si la scène observée est partiellement connue, les approches utilisées vont être très différentes. Nous ne nous intéresserons qu'aux méthodes visuelles que nous pouvons classer dans le schéma de la FIGURE 1. Dans un premier temps, nous pouvons distinguer, les méthodes dites sans structure dans lesquelles aucun objet de l'environnement ou de la scène n'est connu initialement, des méthodes avec structure dans lesquelles certains objets de la scène sont connus et parfois localisables dans un repère monde.

Les méthodes dites sans structure opèrent donc sans aucune connaissance initiale des objets présents dans la scène et de leurs positions. Ces méthodes nécessitent donc de calculer simultanément la pose avec la position d'objets 3D qui serviront de points de repère. Lorsque la cartographie de l'environnement doit être faite en temps réel, nous

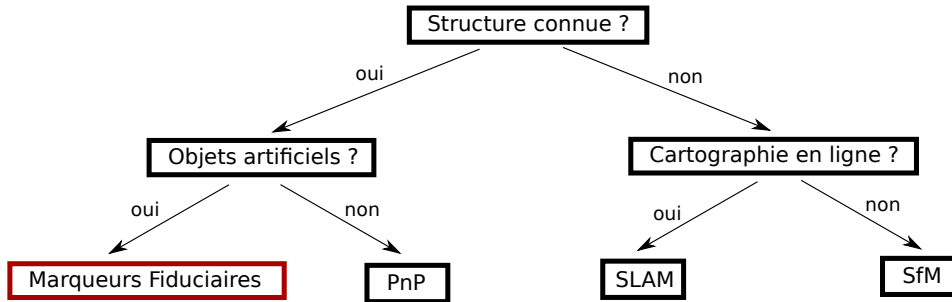


FIGURE 1 – Schéma très simplifié des catégories de méthodes de localisation visuelle.

parlons souvent de SLAM, *Simultaneous Localization and Mapping* ou cartographie et localisation simultanées en français. Le principe consiste à trouver des correspondances de points entre plusieurs images provenant de différents points de vue afin de reconstruire la scène (ou la structure) tout en estimant les poses relatives des caméras (le mouvement). Nous pouvons distinguer ces méthodes des méthodes dites de SfM globale, *Structure from Motion* ou structure acquise à partir du mouvement où la cartographie peut être réalisée en amont, hors ligne, en utilisant plusieurs images de la scène. Pour être plus précis, il faudrait mettre à part les méthodes de SfM incrémentales qui s'apparentent au SLAM, du fait que les images sont traitées une à une de façon incrémentale. Une classification plus détaillée de ces méthodes est présentée dans [82].

En revanche, lorsque les objets dans la scène sont connus le problème revient à l'estimation de la pose à partir de correspondances entre des objets 3D et leurs images 2D, par projection. L'intérêt de ces méthodes est qu'elles peuvent fonctionner sans avoir besoin d'effectuer le suivi de la trajectoire. L'estimation de la pose peut être faite à chaque acquisition. Lorsque les objets utilisés sont des points, le problème est appelé *PnP*, *Perspective n Points*. Ces méthodes nécessitent de connaître les positions de ces points dans l'espace (potentiellement obtenues en amont avec une méthode sans structure) et d'identifier leurs images sur une nouvelle vue. Les correspondances entre la position des points en 3D et leurs images par la projection permettent de calculer la pose de la caméra de prise de vue. La pose peut également être estimée à partir d'objets construits spécifiquement pour cette tâche. Ces objets s'appellent des marqueurs fiduciaux, ces marqueurs sont ainsi placés et fixés dans la scène. De plus, ils possèdent des caractéristiques distinctives qui permettent de les détecter facilement contrairement aux objets naturels tels que les points d'intérêts. Enfin, à partir de l'apparence de leur image par une caméra, il est possible de calculer leur pose dans l'espace. La correspondance entre la forme du marqueur et son image par une caméra contraint la pose de sorte qu'il est possible de calculer la pose du marqueur par rapport à la caméra, c'est-à-dire de retrouver les coordonnées du marqueur dans le repère de la caméra. C'est cette dernière catégorie de méthodes qui va nous intéresser au cours de cette thèse.

Ainsi, nous pouvons distinguer quatre sous-problèmes aux méthodes s'appuyant sur les marqueurs :

- La détection de la projection du marqueur dans l'image ;
- La localisation ou le calcul de la pose du marqueur dans le repère de la caméra ;
- L'identification du marqueur observé dans la scène ;
- La localisation de la caméra dans le repère de la scène.

Pour répondre à ces sous-problèmes, les marqueurs fiduciaux prennent des formes

spécifiques dont leurs images par projection contraignent la pose de celui-ci. Les plus courants sont composés d'un contour carré ou circulaire et, ainsi, leur image par une projection perspective est respectivement un quadrilatère ou une ellipse. Nous allons spécifiquement nous intéresser à ces derniers, les marqueurs circulaires, pour des raisons que nous détaillerons par la suite. Le problème principal que nous allons résoudre concerne le second sous-problème : le calcul de la pose du marqueur à partir de la vue de son contour circulaire. Le calcul de la pose des marqueurs circulaires se fait en utilisant l'image des contours des cercles le définissant, qui correspond à des ellipses à détecter dans la photographie.

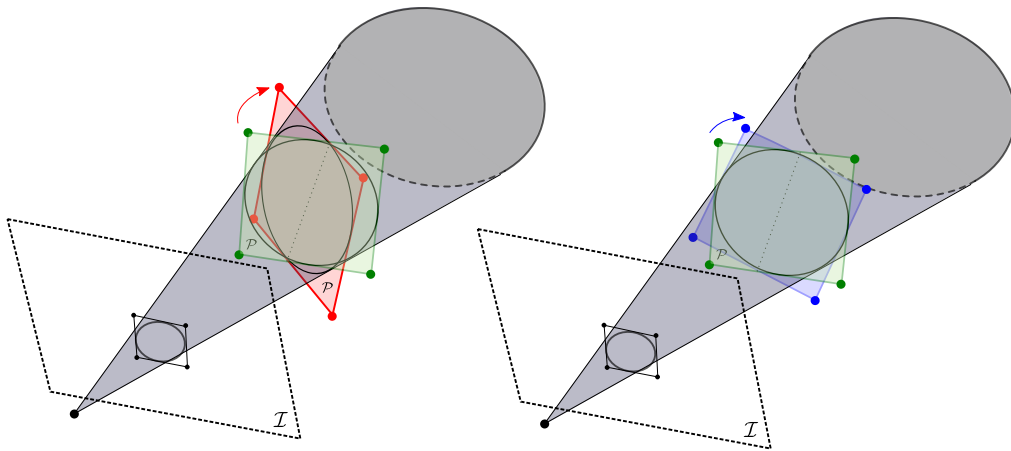


FIGURE 2 – Illustration de l'ambiguïté – Nous présentons l'ambiguïté double sur la pose du plan (à gauche) et l'ambiguïté de rotation autour de l'axe du cercle (à droite).

En réalité, cette seule information n'est pas suffisante pour retrouver la pose d'un marqueur, mais c'est la seule disponible dans certains cas d'utilisation. Lorsque seule l'image du cercle contour est disponible, plusieurs solutions sont possibles et la FIGURE 2 illustre les ambiguïtés possibles donnant la même image du contour.

Nous avons finalement trois sous-problèmes que nous allons étudier dans ces travaux :

- Un problème sous-contraint qui nécessite l'ajout d'un second cercle ou de l'information de calibrage ;
- Une double ambiguïté sur la pose, et ceci même dans le cas calibré ;
- Une rotation inconnue autour de l'axe du cercle due à la symétrie du marqueur.

En conclusion, voici la structure du document qui permet de présenter le travail réalisé :

- Chapitre 1 : Dans ce chapitre introductif, nous présentons les marqueurs fiduciaires et les problématiques associées à savoir : la détection, la localisation et l'identification. Nous présenterons les manières de prendre en compte ces difficultés dans la littérature. Nous concluons avec la présentation du marqueur circulaire avec lequel nous travaillons au cours de ce manuscrit.
- Chapitre 2 : Le second chapitre se concentre sur les méthodes de calcul de l'homographie dans le cadre des marqueurs circulaires. Cette transformation est importante, car ses paramètres donnent des contraintes sur la pose du marqueur. De plus, la rectification obtenue permet d'identifier le marqueur détecté. Nous classons les

méthodes de calcul selon les primitives utilisées (cercles) et leurs dispositions. Enfin, nous étudions une comparaison entre deux méthodes de référence, une méthode qui utilise un marqueur composé de deux cercles concentriques et une méthode qui utilise un seul cercle, avec une caméra calibrée. Les résultats que nous obtenons nous amènent à considérer une nouvelle piste d'étude : Peut-on utiliser seulement le cercle contour et approximer les paramètres intrinsèques de la caméra ?

- Chapitre 3 : Dans le troisième chapitre, nous répondons à cette question. Après nous être placés dans le cas où seulement le cercle contour est détecté, nous proposons une méthode qui s'appuie sur une nouvelle formulation et des *a priori* sur le calibrage de la caméra. Cette méthode permet de générer des candidats pour l'homographie. Nous proposons ensuite d'évaluer ces candidats à l'aide de critères photométriques construits sur un modèle d'apparence du marqueur.
- Chapitre 4 : Lorsque nous cherchons à estimer la pose d'un cercle uniquement à partir de son image sur une vue, il existe une ambiguïté qu'il est impossible de résoudre sans l'ajout d'information supplémentaire. Le quatrième chapitre nous permet d'étudier de façon plus détaillée la formulation proposée au chapitre 3 et d'étudier cette ambiguïté. Nous y développons notamment une condition à partir d'un paramétrage minimal de la pose du cercle qui, lorsqu'elle peut être évaluée, permet de sélectionner la bonne solution algébrique au problème du calcul de la pose à partir du contour.
- Chapitre 5 : Dans le dernier chapitre, nous cherchons à identifier et localiser un marqueur circulaire en utilisant conjointement son contour et des points détectés dans son environnement. Dans ce dernier chapitre, nous résolvons ainsi l'estimation de la rotation autour de l'axe du marqueur et l'ambiguïté de pose tout en identifiant un marqueur dans une bibliothèque d'images de référence.

Chapitre 1

Introduction aux marqueurs fiduciaires

Sommaire

1.1	Définition du marqueur fiduciaire	8
1.2	Différents types de marqueurs plans	10
1.3	Détection d'un marqueur dans l'image	12
1.3.1	Prétraitement de l'image	12
1.3.2	Détection du contour du marqueur	13
1.3.3	Validation des candidats détectés	16
1.4	Estimation géométrique de la pose d'un marqueur	17
1.4.1	Estimation précise des paramètres du contour	17
1.4.2	Estimation de l'homographie	17
1.4.3	Calcul de la pose dans le cadre des marqueurs fiduciaires	18
1.5	Reconnaissance et identification d'un marqueur	18
1.6	Conclusion et présentation d'un marqueur spécifique, le bleam	19

1.1 Définition du marqueur fiduciaire

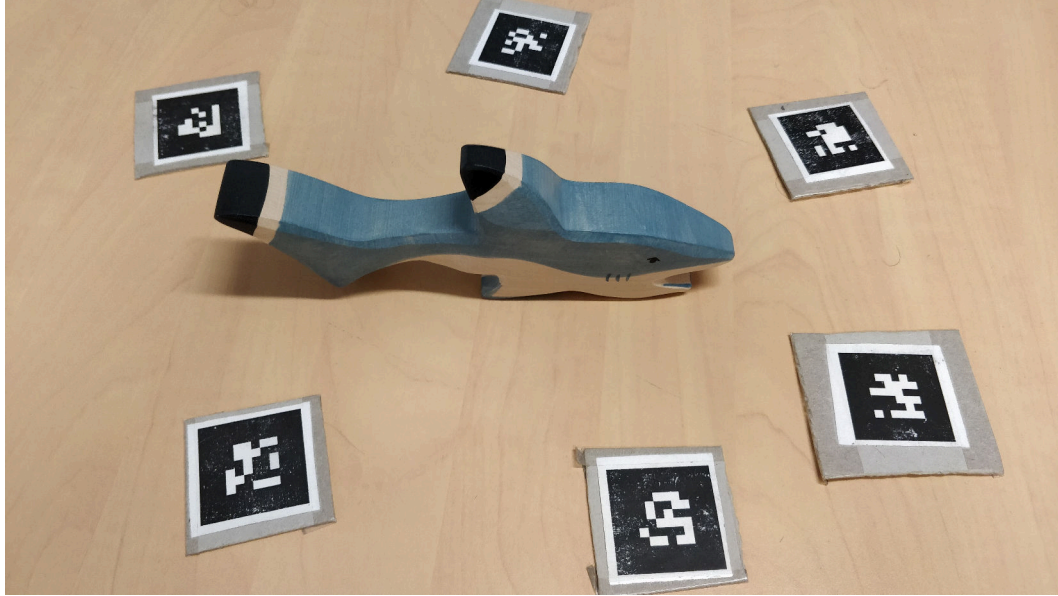


FIGURE 1.1 – Exemple d’un système de marqueur utilisé pour la reconstruction d’objets

Les marqueurs fiduciaux sont, de façon générale, des objets artificiels insérés dans une scène réelle afin de simplifier les problèmes de reconnaissance d’objets ou d’estimation de pose. Un système de marqueurs est composé d’un ensemble de marqueurs associés chacun à un identifiant unique et des algorithmes permettant leur détection ainsi que la lecture de leurs identifiants. La FIGURE 1.1 montre un exemple de système de marqueur utilisé pour reconstruire un objet en 3D, les marqueurs utilisés sont des ARTag (FIGURE 1.2b). Il existe de nombreux systèmes de marqueurs et il est difficile d’en faire une présentation exhaustive, c’est pourquoi, nous préférons ici reprendre la manière de décrire les marqueurs introduite par [20]. Dans cet article, il est proposé de distinguer onze propriétés que peuvent avoir les marqueurs fiduciaux afin de fournir un système fiable de marqueurs. Voici ces onze critères :

1. Avoir une taille minimale du marqueur dans l’image (en pixel) ;
2. Être robuste aux occultations ;
3. Être robuste aux changements de luminosité ;
4. Être indépendant des paramètres photométriques du support du marqueur et de l’environnement ;
5. Limiter le taux de faux positifs (fausses détections) ;
6. Limiter le taux de faux négatifs (manque de détections) ;
7. Être performant par rapport au temps d’exécution ;
8. Présenter un ensemble de primitives suffisantes pour le calcul de la pose ;
9. Éviter l’instabilité de calcul de la pose du marqueur ;
10. Augmenter la taille de la bibliothèque de marqueurs ;
11. Limiter la confusion inter marqueurs.

Les six premiers critères concernent la capacité du marqueur à être facilement détecté dans une image. Cela implique, en particulier, que le marqueur ait la capacité de se distinguer du reste de la scène, quelle que soit la scène. Ainsi, la robustesse des marqueurs aux conditions de prises de vues (éclairage, occultation, faible résolution) est un critère important. C’est souvent la raison pour laquelle, dans la littérature, des formes de marqueurs spécifiques sont utilisées.

Les cinq propriétés suivantes concernent les critères d’évaluation de la qualité de la détection. Algorithmiquement, la détection est, en général, la partie qui nécessite le plus de calculs et donc, en conséquence, elle demande des temps d’exécution élevés [15]. La position dans l’espace du marqueur par rapport à une caméra perspective doit pouvoir être entièrement contrainte par son apparence. Ce critère confère au marqueur la capacité d’être utilisé pour de nombreuses applications comme la réalité augmentée, la localisation de robots ou drones, etc. Les critères (8) et (9) concernent le calcul de la pose du marqueur à partir de son image détectée au travers d’une ou de plusieurs vues. Le premier critère signifie qu’il est possible de calculer la pose du marqueur à partir de son observation dans une image. Le second garantit que cette estimation est précise et consistante entre plusieurs vues du marqueur au cours du temps. Nous reviendrons particulièrement sur ces critères et la mise en place d’un système de marqueurs et d’algorithmes qui garantissent ces propriétés dans le reste du manuscrit. L’intérêt de pouvoir calculer la pose du marqueur est double, elle permet d’une part de rectifier l’image du marqueur dans son plan de support afin de simplifier la lecture de son identifiant. D’autre part elle permet de calculer la pose relative du marqueur par rapport aux caméras pour appliquer de la réalité augmentée par exemple. Enfin les deux derniers critères concernent l’identification du marqueur dans une bibliothèque. Ces deux derniers critères ne sont pas toujours compatibles et il s’agit généralement d’un compromis. En effet plus la taille de la bibliothèque est grande, plus deux marqueurs aux identifiants différents vont être difficiles à distinguer entre eux. Beaucoup de systèmes de marqueurs proposent ainsi des modèles ajustables en fonction du nombre de marqueurs distincts souhaités dans l’application [4, 20].

Les marqueurs fiduciaux majoritairement les plus étudiés en vision par ordinateur sont les marqueurs plans. Un marqueur plan est un marqueur associé à un plan de support tel qu’il est possible de calculer (sans ambiguïté), à partir d’une nouvelle vue, les propriétés de la transformation qui change l’image du marqueur sur son plan de support en l’image du marqueur sur cette nouvelle vue. Dans le cadre de la projection perspective d’un marqueur plan, cette transformation est une homographie (cf. annexe A.4). Le calcul préliminaire de cette homographie, dans le cas où la caméra est calibrée, permet de calculer la pose du marqueur. Ceci simplifie le critère (8) qui revient alors au calcul de l’homographie qui donne à la fois la rectification de l’image du marqueur dans son plan de support et les paramètres de la pose du marqueur. Les marqueurs plans constituent en fait la forme géométrique la plus simple suffisante pour calculer la pose de l’objet à partir d’une modélisation perspective de la projection.

On peut tout de même noter que quelques marqueurs 3D ont été proposés mais, en général, ils se composent d’une ou plusieurs parties planes et d’une autre composante servant à identifier d’autres caractéristiques. Un exemple de ce type de marqueur est celui proposé par Kanbara *et al.* [38]. Le marqueur permet de calculer les caractéristiques photométriques de l’environnement à partir d’un miroir sphérique placé sur un marqueur carré. Ces caractéristiques permettent de générer des éléments virtuels dans une application de réalité augmentée aux conditions d’éclairage plus réalistes, en utilisant la cartographie de réflexion obtenue grâce au miroir.

Dans la section suivante, nous allons ainsi décrire les marqueurs plans les plus connus et les plus utilisés.

1.2 Différents types de marqueurs plans

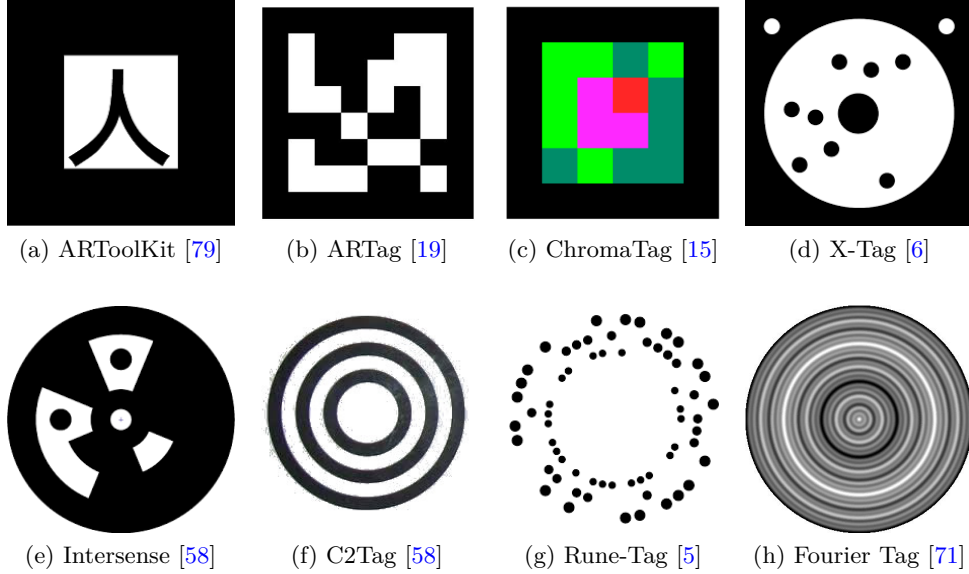


FIGURE 1.2 – Différents types de marqueurs fiduciaux.

Les marqueurs peuvent prendre des formes diverses et variées en fonction des caractéristiques ou propriétés, comme celles décrites dans la section 1.1, qui sont prioritaires pour l'application visée. Nous pouvons distinguer deux grandes classes de marqueurs selon la forme de leur contour externe : les marqueurs à contour carré ou circulaire. Les marqueurs à contour carré présentent l'avantage de contraindre l'homographie avec la position de l'image de leurs quatre coins bien qu'il existe 4 ambiguïtés à cause de la symétrie. Cependant ils sont en conséquence plus sensibles à l'occultation d'un de ces coins dans l'image et également plus sensibles à une erreur de mesure qui peut alors affecter la précision du calcul de l'homographie. Pour le cas des marqueurs à contours circulaires, la seule image du contour externe (une ellipse) ne contraint pas l'homographie et ceci même dans le cas où la caméra serait calibrée comme nous le verrons en détail dans le chapitre 2. Les marqueurs disposent donc en général d'une autre primitive détectable dans l'image ou d'un second contour circulaire. Au-delà de l'intérêt esthétique pour les marqueurs circulaires, ces derniers sont souvent préférés pour leur résilience à une occultation partielle de l'image de leur contour. En effet une ellipse peut être estimée à partir de seulement 5 points quelconques de son contour ainsi un marqueur partiellement occulté peut toujours être correctement localisé. Localiser un marqueur à contour circulaire consiste donc tout d'abord à détecter l'ellipse, l'image de son contour.

Concernant les marqueurs carrés, le code à l'intérieur peut soit être encodé par un motif intérieur quelconque comme c'est le cas du marqueur d'ARToolKit (FIGURE 1.2a) ou bien être composé d'une matrice binaire de données en noir et blanc comme l'ARTag (FIGURE 1.2b) ou en couleur comme le ChromaTag (FIGURE 1.2c). Le marqueur X-tag (FIGURE 1.2d) contient un contour interne circulaire pour augmenter la précision de localisation et un système de points aléatoires préalablement appris dans une

bibliothèque.

Concernant les marqueurs circulaires, de l'information peut être stockée dans la forme des secteurs comme pour le marqueur Intersense (FIGURE 1.2e), dans la taille des anneaux comme pour le C2Tag (FIGURE 1.2f), ou encore dans la présence ou non de points dans le contour du marqueur pour le Rune-Tag (FIGURE 1.2g). Si la plupart des marqueurs ont une distance limite d'opérabilité en dessous de laquelle ils peuvent être entièrement reconnus, certains marqueurs comme le Fourier Tag (FIGURE 1.2h) présente un système d'identification qui dépend de la distance à laquelle ils sont observés. Plus il est observé de près, plus l'information décodée sur ce marqueur sera précise. Son identifiant est encodé dans la transformée de Fourier du signal, ainsi les informations précises sont stockées dans les hautes fréquences spatiales de l'image et les informations moins précises dans les basses fréquences.

Dans ce manuscrit, nous avons fait le choix de ne pas présenter en détail chacun de ces algorithmes, mais plutôt de présenter un algorithme assez générique de la manière de procéder pour ces différents marqueurs. En effet, bien que les marqueurs puissent être très différents, les algorithmes qui les utilisent présentent des étapes similaires, la FIGURE 1.3 résume ces étapes.

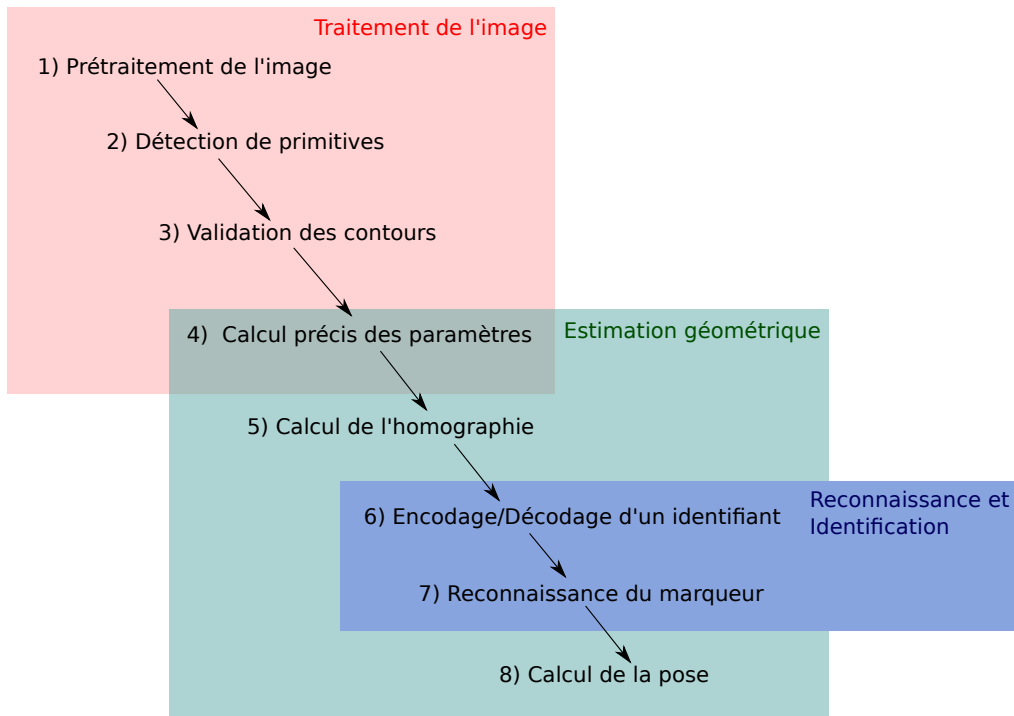


FIGURE 1.3 – Étapes de reconnaissance des marqueurs fiduciaux

Suite à cette introduction sur les marqueurs fiduciaux, nous pouvons nous demander quelle est la forme de marqueurs que nous pouvons privilégier. Ce choix dépend essentiellement des propriétés recherchées et du critère qui nous intéresse le plus parmi les 11 critères identifiés. À la fin de ce chapitre, nous présenterons ainsi le marqueur que nous avons décidé d'étudier.

Dans nos travaux, nous allons étudier en particulier les étapes 5, 7 et 8, c'est-à-dire le calcul de l'homographie, l'identification du marqueur et le calcul de la pose du marqueur. Toutefois, les autres étapes sont cruciales, car elles influencent fortement la qualité de la

pose complète retrouvée. Nous proposons donc d'étudier un peu plus en détail la façon dont ces étapes ont été perfectionnées dans la suite de cet état de l'art.

1.3 Détection d'un marqueur dans l'image

La détection d'un marqueur dans l'image consiste à retrouver dans l'image la région correspondant à l'image d'un marqueur possiblement présent. Il est important également de souligner qu'il y a deux aspects : la détection de la primitive et éventuellement son suivi. Le suivi consiste à suivre l'évolution de la position d'un marqueur précédemment détecté. On parle généralement de suivi par détection ou de détection par suivi. La première méthode consiste à effectuer des détections sans utiliser les *a priori* des précédentes détections, ainsi le suivi consiste simplement à fournir les positions successives des marqueurs détectés. La seconde méthode consiste à suivre le marqueur en utilisant un modèle qui interprète et prédit la trajectoire (chemin tracé par les positions successives) du marqueur ce qui facilite la détection du marqueur dans une nouvelle acquisition. Si certains systèmes de marqueurs utilisent parfois cette dernière méthode, dans nos travaux nous supposons que les détections sont indépendantes. En effet le risque de perdre le suivi est important notamment à cause d'occultations ou d'une vitesse de mouvement élevée par rapport à la fréquence d'acquisition, ce qui introduit du flou dans l'image.

1.3.1 Prétraitement de l'image

Une première étape consiste à localiser les régions de l'image dans lesquelles un marqueur peut être présent. Pour cette étape, la forme du marqueur (généralement carré ou circulaire) a peu d'importance, il s'agit plutôt de détecter les régions connexes de l'image.

Lorsque les marqueurs représentent un polygone ou un disque noir sur fond blanc, ils peuvent être simplement détectés par seuillage sur l'intensité comme proposé par Rekimoto *et al.* [69]. Toutefois, il s'avère que l'intensité de l'image peut varier localement du fait des variations d'éclairage. Ainsi, certains auteurs proposent d'appliquer en amont un filtrage homomorphique pour diminuer les effets liés aux variations d'éclairage dans l'image et ainsi augmenter le contraste [58, 63] quand cela est nécessaire. Ces méthodes par seuillage reposent également sur le choix d'un paramètre global et ce paramétrage supplémentaire peut s'avérer difficile à ajuster. Des seuillages automatiques ont donc été proposés pour remédier à ce problème. Notamment, les auteurs de [79] proposent d'utiliser un seuil utilisant la médiane de l'intensité des marqueurs observés précédemment ou de le sélectionner aléatoirement en cas d'initialisation ou de perte de suivi. D'autres auteurs proposent des méthodes de seuillage local consistant à exploiter les intensités maximales et minimales détectées aux alentours de chaque pixel [80].

Une autre classe de méthodes consiste à détecter directement les contours des marqueurs dans l'image en travaillant sur le gradient de l'image. Il s'agit de calculer les différences d'intensité (qui sont liées au calcul du gradient) et ainsi, de faire ressortir les régions de fort contraste, en supposant qu'elles correspondent à des contours. Ces méthodes se sont révélées plus efficaces pour détecter le marqueur lorsqu'il y a d'importantes variations d'intensité. Dans [58], les auteurs utilisent un filtre simple et classique, celui de Sobel. Celui-ci est facile à mettre en œuvre, mais il est peu robuste aux difficultés, comme par exemple, un contour faiblement contrasté ou un contour occulté. Ainsi, il est classique d'utiliser le filtre de Canny afin de rendre robuste ce type de filtre [19, 63, 7].

En effet, cet opérateur tente de relier les parties de contours qui peuvent être déconnectées à cause de la présence d'une occultation. Pour cela, il s'appuie sur un seuillage par hystérésis (utilisation de deux seuils) pour accepter ou rejeter des points contours. Ainsi, plus précisément, un point est :

- rejeté si l'intensité de son gradient est inférieure au premier seuil ;
- conservé si elle est supérieure au second seuil ;
- conservé si sa valeur est entre les deux seuils et que ce point est connecté à un point conservé, rejeté sinon.

Ce processus est réalisé de manière itérative jusqu'à ce qu'il n'y ait plus de changements. En conséquence, ce filtre est plus coûteux en temps de calcul. Sur la FIGURE 1.4, nous pouvons observer les différents filtrages et leur efficacité à détecter correctement les contours. Le filtrage binaire (FIGURE 1.4b) par seuillage permet d'obtenir le contour externe sur cet exemple, même en prenant un seuil arbitraire comme la médiane. Le filtrage de Sobel (FIGURE 1.4c) permet d'obtenir les deux premiers cercles. Finalement le filtrage de Canny (FIGURE 1.4d) montre l'intérêt de l'approche qui permet d'obtenir tous les contours.

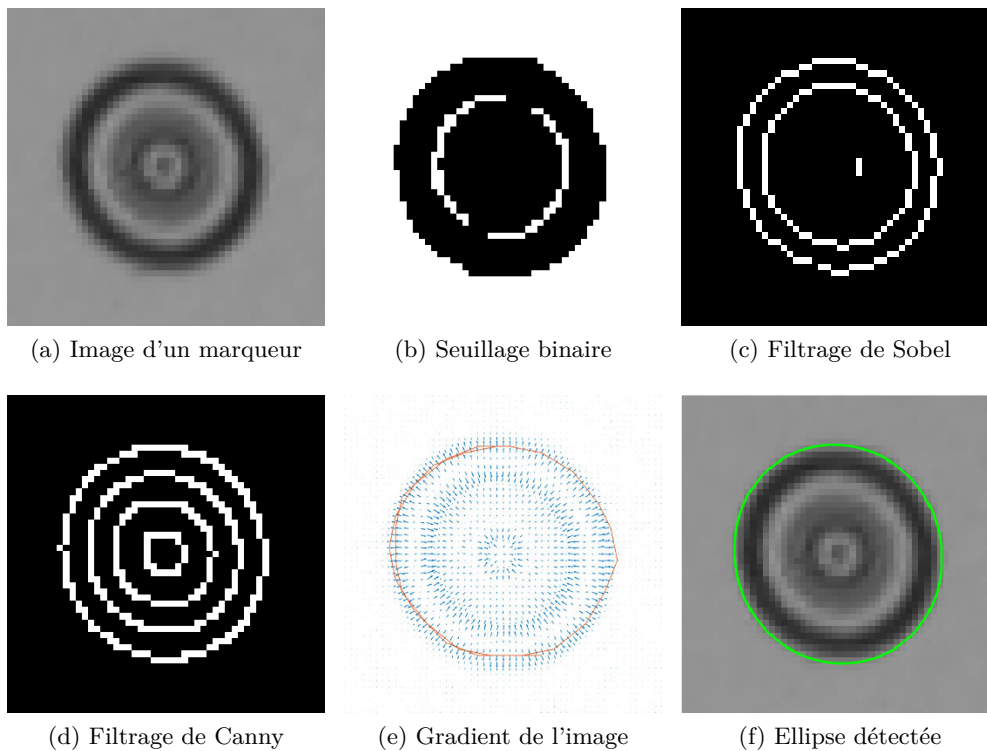


FIGURE 1.4 – Procédé de détection d'un contour circulaire

1.3.2 Détection du contour du marqueur

À partir d'une région ou d'un contour connexe, la forme de l'image du marqueur doit être retrouvée, concrètement un quadrilatère pour les marqueurs carrés et une ellipse pour les marqueurs circulaires. Nous rappelons dans ce paragraphe que nous supposons que nous avons une détection préalable, mais que celle-ci est bruitée (mal positionnée sur le contour), voire incomplète (des parties du contour sont manquantes) et tous les traitements décrits dans la suite visent à corriger ces défauts.

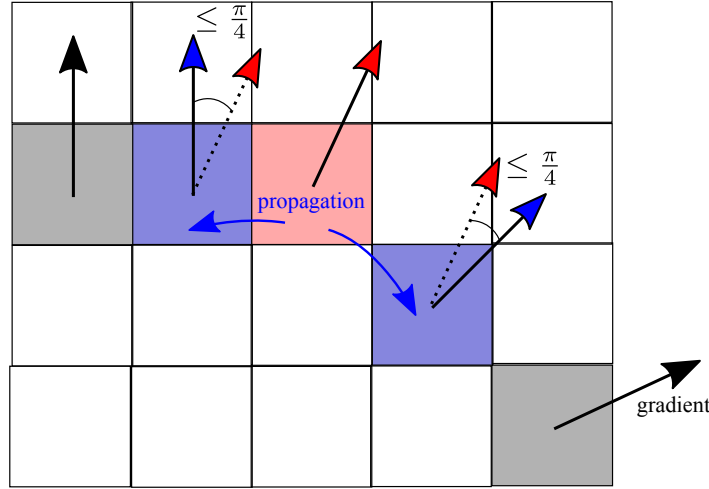


FIGURE 1.5 – Illustration du procédé de propagation pour la détection du contour

Adaptations pour les détections de type quadrilatère. Pour les marqueurs à contour rectangulaire, un quadrilatère doit être ajusté sur les points contours ou sur la région du marqueur. Rekimoto *et al.* proposent d'ajuster sur l'image binaire de la région le quadrilatère [69]. Pour cela ils proposent de minimiser au moindre carré la distance entre le contour de la région et un quadrilatère. Fiala [19] propose d'ajuster séparément les quatre segments du contour détecté et de les joindre pour obtenir le quadrilatère pour atteindre une plus grande précision. Pour ajuster simultanément les quatre segments, Wang *et al.* proposent une méthode permettant d'assimiler ce problème d'ajustement de quatre segments à un problème de partitionnement des points contours en quatre partitions [80]. Ainsi, dans un premier temps, les points contours sont triés en utilisant l'angle autour du centroïde de la région. Des coins candidats sont ensuite sélectionnés en calculant avec une fenêtre glissante sur cette liste triée, un critère permettant d'évaluer l'alignement des points voisins. Les plus « mauvais » points sont alors considérés comme des coins potentiels. L'adéquation de ces quadrangles proposés avec tous les autres points contours est ensuite évaluée afin de déterminer le quadrilatère final.

L'algorithme proposé par [15] consiste à construire un polygone en calculant les points contours à partir de ligne de coupe de l'image. Une ligne de coupe est une image unidimensionnelle construite à partir d'une droite de l'image. Dans l'algorithme décrit par Degol *et al.*, celles-ci sont initialement générées par des droites tirées aléatoirement. Sur une ligne de coupe, les points contours peuvent être trouvés en recherchant les transitions colorées rouges/vertes (voir le marqueur de la FIGURE 1.2c). Ce polygone est agrandi itérativement en ajoutant des points à partir de nouvelles coupes. L'algorithme converge finalement vers un quadrilatère dont les côtés sont les quatre plus longs segments. Cette méthode permet d'améliorer la rapidité en évitant d'avoir à extraire le contour en entier.

Adaptations pour les détections de type ellipse. Pour la détection d'ellipse, le plus simple et le plus intuitif, dans la littérature, est de prendre en compte la courbure du contour. En effet, contrairement au quadrilatère, le contour d'une ellipse possède une courbure continue. Ainsi, les deux gradients de points consécutifs du contour possèdent une direction proche, comme nous pouvons l'observer sur la FIGURE 1.4e. La détection

du contour peut ainsi se faire en liant de petits segments du contour les uns aux autres, en considérant que la direction de leur gradient doit être continue [65, 22]. Ainsi, pour Fornaciari *et al.*, les étapes suivantes sont réalisées [22] :

- génération de germes à partir de l'extraction des gradients (filtre de Sobel) ;
- propagation dans les deux directions perpendiculaires au gradient sur les points contours voisins pour former un arc ;
- agrégation d'un point contour voisin à l'arc si la direction de son gradient ne s'éloigne pas de plus d'un octant dans le sens de la propagation.

La génération de germes se fait en recherchant des pixels avec une norme du gradient élevé et présentant les caractéristiques d'un contour de courbe ; c'est-à-dire que les pixels voisins présentent également une norme de gradient élevée et une direction proche (moins d'un octant). La FIGURE 1.5 illustre le principe de propagation et la FIGURE 1.4e montre son application sur notre image d'exemple. Nous pouvons observer que plusieurs germes ont été détectés, ce qui donne naissance à plusieurs arcs. **Cette méthode est celle que nous avons choisie dans la thèse pour détecter les contours elliptiques dans les images.** Ce choix est justifié par l'efficacité de l'algorithme dans le cas de la détection d'ellipse, celui-ci est en effet efficace pour détecter des arcs elliptiques malgré des ellipses de très petites dimensions sans demander beaucoup de temps d'exécution. Patraucean *et al.* proposent une approche assez similaire mais se basant sur la construction de segments [65] : ces segments sont construits en assemblant les points contours voisins dans la direction opposée de leur gradient de sorte que leurs gradients partagent une même direction. Ces micro segments peuvent ensuite être assemblés dans une même courbe si son gradient moyen ne varie pas trop entre deux segments consécutifs.

Lorsque les ellipses à détecter possèdent une grande excentricité, il se peut que deux arcs ne soient pas correctement reliés. Ainsi, une méthode d'agrégation d'arcs est nécessaire pour estimer une nouvelle ellipse plus complète à partir d'arcs de courbes disjoints. Ré-estimer une ellipse pour chaque paire d'arcs candidats est très coûteux en temps de calcul. Ainsi, Nguyen *et al.* proposent d'évaluer des critères sur les angles relatifs et la distance entre deux arcs pour rejeter rapidement des arcs candidats incompatibles [59]. Cependant pour des raisons de coût en temps de calcul, nous n'utilisons pas cette méthode dans nos applications. En effet, la méthode que nous utilisons dans nos expérimentations se concentre sur la détection d'un contour en entier en continuant la propagation jusqu'à revenir au point initial. Cette méthode évite ainsi d'obtenir des segments d'arcs non connectés, mais nécessite également d'être moins exigeant sur la norme du gradient dans la propagation à cause du risque de ne pas détecter le contour.

Une fois le contour elliptique obtenu, une ellipse peut être ajustée sur ce contour. L'ajustement d'une ellipse consiste à trouver l'ellipse la plus proche en distance euclidienne des points du contour. Ce problème n'est pas trivial car le calcul de la distance d'un point à une ellipse nécessite la résolution d'un système polynomial de degré quatre (équation quartique). En général, l'estimation se fait sur des distances algébriques ou des mesures qui approximent la distance euclidienne. Les premières méthodes d'estimation de l'ellipse très populaires et très utilisées en raison de leur simplicité et de leur efficacité sont celles de Fitzgibbon [21] et Halir [28]. Ces méthodes d'ajustement très rapides reposent sur la résolution du problème aux moindres carrés linéaires avec des contraintes elliptiques. Le problème est linéarisé dans un premier temps en transformant l'espace des données dans un espace de dimension plus grand, soit $\mathbf{X}_i = [x_i^2 \ x_i y_i \ y_i^2 \ x_i \ y_i \ 1]^\top \in \mathbb{R}^6$. Le problème revient alors à trouver un vecteur de paramètres $\theta = [a \ b \ c \ d \ e \ f] \in \mathbb{R}^6$ tel que $\mathbf{X}^\top \theta = 0$. L'ajout d'une contrainte quadratique permet de garantir que la conique ainsi estimée est bien une ellipse et non

une conique quelconque : $4ac - b^2 = 1$. La résolution est détaillée dans l'annexe B.1. La FIGURE 1.4f illustre le résultat obtenu sur le contour de notre marqueur circulaire. Cette méthode pose des problèmes de biais importants puisque la minimisation se fait dans un espace différent de celui des données initiales du problème mais elle reste une approximation correcte utilisée par quelques marqueurs circulaires [24, 63], en particulier, pour la rapidité de l'approche. Pour d'autres marqueurs, cette méthode n'est utilisée que pour vérifier que les contours obtenus sont cohérents avec une ellipse et, ensuite, une estimation plus précise est utilisée. Nous détaillons ces approches dans l'annexe B.

Nous n'avons pas cité toutes les méthodes existantes de manière exhaustive, puisque ce n'est pas le cœur du travail développé dans ce manuscrit. Toutefois, il paraît important d'en décrire quelques-unes des plus utilisées pour bien comprendre les difficultés liées à l'estimation du contour d'un marqueur. Notamment, nous mettons en évidence la présence d'erreurs plus ou moins importantes sur les points qui constituent les contours et qui peuvent induire des erreurs non négligeables sur l'estimation de la pose.

1.3.3 Validation des candidats détectés

Un des problèmes majeurs qui limite l'intérêt de l'utilisation d'un marqueur est le taux important de faux positifs qui peuvent être détectés. Or, une détection erronée, si elle est conservée dans la chaîne de traitement, aura un impact conséquent sur les performances de l'algorithme. Des approches permettent donc de rejeter rapidement les formes qui ne sont pas cohérentes vis-à-vis du modèle d'apparence d'un marqueur. Par exemple, il est possible d'exploiter des *a priori* sur les marqueurs en ajoutant des hypothèses sur la taille et le rapport hauteur/largeur des marqueurs, comme dans [69]. Des tests sur la photométrie peuvent également être ajoutés : soit uniquement sur le niveau de gris [58], soit sur un canal de couleur spécifique, comme le canal a^* de l'espace colorimétrique L^*a^*b CIE [15]. Une autre possibilité est d'utiliser une approche statistique *a contrario* pour déterminer si un contour détecté correspond plus probablement à du bruit, à une droite ou à une courbe elliptique [65]. Pour cela, la structure de l'image est analysée pour mesurer la probabilité qu'un arc elliptique apparaisse aléatoirement. Par exemple la proportion de gradients orthogonaux à la tangente de l'ellipse est mesurée. Cette mesure de la vraisemblance de l'apparition d'une telle structure par rapport à une image aléatoire permet de mesurer la probabilité de fausse détection. Pour les marqueurs circulaires, Lightbody *et al.* proposent d'évaluer si l'aire des pixels de la forme détectée correspond à l'aire de l'ellipse estimée [46]. Bergamasco *et al.* utilisent la forme particulière de leur marqueur [5]. En effet, celui-ci est composé de nombreux cercles qui sont positionnés en suivant le contour (non visible) d'un plus grand cercle, cf. FIGURE 1.2g. Ainsi, les petits cercles sont détectés dans un premier temps et ils sont utilisés pour estimer les paramètres de l'ellipse, image du grand cercle auquel ils appartiennent. Chaque petit cercle vote pour le grand cercle auquel il appartient, le plus probablement. Ainsi, le résultat de ce vote permet de valider le contour de l'image du grand cercle en vérifiant que l'agencement est correct. Le marqueur de Benini *et al.* [3] présente aussi un agencement particulier de cercles de rayons différents. Le marqueur est constitué d'un anneau externe contenant 5 disques noirs, 1 au centre et 4 sur chacun de ses côtés avec un cercle de rayon plus grand pour contraindre l'orientation. Une fois les contours elliptiques détectés dans l'image, une vérification peut être effectuée sur leur agencement pour qu'il soit cohérent avec l'apparence générale du marqueur.

1.4 Estimation géométrique de la pose d'un marqueur

1.4.1 Estimation précise des paramètres du contour

Pour les marqueurs circulaires, la précision des paramètres de l'ellipse estimée sur le contour est essentielle, c'est en effet ces paramètres qui serviront à rectifier l'image et garantiront l'efficacité du marqueur. La méthode linéaire présentée plus tôt a le défaut de transformer l'espace de résolution du problème au lieu de directement établir une minimisation sur l'espace des données, ceci introduit un biais sur l'estimation. Afin de réduire un peu ce biais, il est possible de normaliser les données initiales de sorte que le barycentre des points soit à l'origine et que leur matrice de covariance ait deux valeurs propres proches de 1 comme proposé par Hartley [30].

Pour une estimation plus précise, d'autres méthodes ont été développées. Nous pouvons distinguer les méthodes reposant sur des approximations de la distance utilisée, que nous présentons dans l'annexe B.2, des méthodes reposant sur une approche statistique comme le maximum de vraisemblance. Cette dernière catégorie s'appuie sur la connaissance d'*a priori* sur l'incertitude des points utilisés pour l'estimation. Ainsi, le problème est assimilé à l'estimation de l'ellipse qui maximise la vraisemblance des points observés. Nous présentons en annexe B.3 ces méthodes.

1.4.2 Estimation de l'homographie

Pour que le marqueur puisse être lu malgré la déformation perspective liée à la projection, il est nécessaire d'estimer les paramètres de la transformation projective. Une fois cette transformation connue il est possible de rectifier l'image entièrement ou de connaître la position des points encodant l'identifiant du marqueur. La projection du plan de support du marqueur par une caméra peut être entièrement paramétrée par une matrice d'homographie, que nous noterons H . Une définition plus complète est présentée en annexe A.4. Les termes de cette matrice sont estimés grâce aux contraintes apportées par la correspondance entre le modèle d'apparence du marqueur et l'observation sur l'image. Plus précisément, nous avons besoin de primitives appariées. Nous appelons primitive un objet géométrique. En général, il s'agit d'un point, d'un segment, d'un polygone ou d'un cercle dont l'image par la projection apporte une contrainte sur l'homographie. L'ensemble de ces primitives couplé aux informations photométriques (couleurs des pixels du marqueur) constitue le modèle d'apparence du marqueur. Nous pouvons distinguer les méthodes en fonction des différents types de primitives utilisées.

Le problème est donc la recherche de l'homographie \mathcal{H} , paramétrée par la matrice H telle que les équations suivantes soient respectées.

$$\mathcal{H}(P_i) = P'_i \text{ avec } P_i \text{ et } P'_i \text{ respectivement la primitive et son image} \quad (1.1)$$

Ce type d'équation qui lie les correspondances entre la primitive et son image fournit des contraintes sur l'homographie et réduit ses degrés de liberté. En effet, nous rappelons qu'une homographie possède huit degrés de liberté et qu'il est donc possible de l'estimer quand l'ensemble d'équations formé par 1.1 fournit plus de huit équations indépendantes. Les méthodes employées varient de façon importante en fonction des primitives disponibles, et celles-ci seront l'objet principal du chapitre 2.

Pour les marqueurs carrés, les quatre coins du quadrilatère sont employés pour contraindre l'homographie, chaque point fournissant deux équations indépendantes sur \mathcal{H} [69, 19, 79, 15]. Il existe cependant quatre homographies ambiguës liées aux axes de symétrie du carré. Cette ambiguïté est résolue grâce à la lecture de l'identifiant.

Pour les marqueurs circulaires, il faut posséder l'image d'au moins deux cercles dans le cas général, chaque correspondance apportant cinq équations indépendantes sur \mathcal{H} . Les C2Tag [7] et les RUNE-Tag [5] utilisent au moins 2 cercles pour calculer l'homographie. Il est important de noter qu'à cause de la symétrie autour du centre du marqueur, il existe une rotation qui ne peut pas être estimée pour les marqueurs circulaires.

Lorsque la caméra est calibrée, les paramètres intrinsèques apportent des contraintes sur l'homographie. L'homographie ne possède ainsi plus que six degrés de liberté qui correspondent aux mêmes degrés de liberté que ceux de la pose de la caméra. Les paramètres intrinsèques sont ainsi utilisés avec la contrainte apportée par une seule primitive circulaire pour le marqueur de Pagani [63], les Whycons [46] ou encore pour les marqueurs proposés par Gosh *et al.* [24]. Dans ce contexte, l'homographie est encore estimée à une rotation près.

1.4.3 Calcul de la pose dans le cadre des marqueurs fiduciaires

Peu d'articles traitent en détail du calcul de la pose pour les marqueurs fiduciaires. Par ailleurs, tous les marqueurs fiduciaires ne permettent pas de calculer la pose de la caméra, comme le FourierTag (FIGURE 1.2h) dont les caractéristiques ne permettent que le décodage.

Toutefois, lorsqu'il est possible de mettre en correspondance suffisamment de points 3D sur le plan du marqueur avec leurs images sur le plan de support, la pose du marqueur peut être estimée, à la condition que les paramètres de la caméra soient connus ou qu'ils puissent être approximés. Ce problème de calcul de la pose à partir de correspondances entre des points 3D et des points 2D est nommé *PnP* (pour *Perspective from n Point*). Il peut être résolu avec au minimum trois correspondances, soit simplement pour les marqueurs carrés ou des marqueurs disposant de l'image de suffisamment de points comme RuneTag [5].

Certains auteurs utilisent la factorisation de l'homographie, en effet si la matrice des paramètres intrinsèques est connue, la pose peut être obtenue en utilisant par exemple la méthode proposée par Sturm [74]. La pose peut être également calculée à partir de correspondances photométriques denses entre le marqueur et son image Edwards *et al.* utilisent la différence photométrique entre un rendu projeté du marqueur et l'image observée [17]. La minimisation de cette différence permet d'obtenir la pose du marqueur.

1.5 Reconnaissance et identification d'un marqueur

Pour identifier un marqueur, il est nécessaire d'utiliser les caractéristiques photométriques, qui sont supposées uniques, et qui le distinguent de l'ensemble des marqueurs de la bibliothèque. En général, une bibliothèque de marqueurs est constituée uniquement des caractéristiques propres aux marqueurs et les caractéristiques utilisées varient selon les systèmes de marqueurs.

Certains systèmes de marqueurs comme celui présenté par Kato *et al.* utilisent un motif quelconque pour chaque marqueur, cf. FIGURE 1.2a. La bibliothèque est donc constituée directement des images des marqueurs. Leur identification est réalisée par une comparaison photométrique directe entre l'image rectifiée du marqueur et les images stockées dans la bibliothèque. En conséquence, cela nécessite de stocker les images des marqueurs et la comparaison peut être lourde en termes de temps de calcul. Cette comparaison photométrique peut également servir à vérifier la pose du marqueur. C'est ce que propose Pagani *et al.* [63] qui utilisent la mesure de corrélation croisée entre

l'image rectifiée du marqueur et celles de sa bibliothèque pour calculer l'orientation du marqueur. De plus, l'image subit différentes rotations avec plusieurs angles et ainsi, ils sélectionnent la rotation qui permet d'obtenir l'image la plus proche dans la bibliothèque. Ainsi, cela permet de retrouver à la fois l'orientation et l'identifiant du marqueur.

D'autres systèmes de marqueurs proposent de générer des points aléatoirement et de stocker uniquement des descripteurs de forme qui décrivent la position relative des points entre eux. C'est le cas du marqueur X-Tag, cf. FIGURE 1.2d, qui utilise l'invariant projectif du birapport anharmonique pour décrire la position relative de quatre points dans un descripteur. L'intérêt de cette mesure est qu'elle est invariante projectivement et que le marqueur peut être lu sans avoir besoin de calculer la rectification de l'image.

En dehors des cas que nous venons de décrire, la plupart des marqueurs utilisent plutôt un codage binaire de l'identifiant. Ceci permet en effet de compresser l'information dans un code de petite taille afin de maximiser la taille de la bibliothèque. La lecture de l'identifiant se fait donc par l'identification en deux classes des parties codantes du marqueur, 0 si l'intensité est faible, 1 si l'intensité est plus élevée. L'encodage peut aussi se faire dans un canal de couleur particulier comme sur le canal b^* pour ChromaTag, cf. FIGURE 1.2c.

En général, les codes binaires ne correspondent pas directement à l'identifiant du marqueur, le plus souvent la redondance est utilisée pour rendre la lecture plus fiable. L'approche la plus utilisée est le contrôle de redondance cyclique [69, 19, 5]. Cette méthode s'appuie sur le calcul d'une valeur de contrôle ajoutée au code. Dans le cas d'une erreur qui ne serait pas trop importante, cette valeur calculée à partir du code initial permet de détecter et de corriger l'erreur. Lightbody *et al.* proposent d'utiliser un autre type d'encodage, celui de Manchester sur le bandeau circulaire de leurs marqueurs [46]. Ce codage a l'intérêt de pouvoir être compris même dans le cas d'un léger décalage de lecture, mais il nécessite de multiplier la taille du code binaire par 2.

Sattar *et al.* détaillent une méthode différente s'appuyant sur la transformée de Fourier pour encoder l'information dans le domaine fréquentiel, cf. FIGURE 1.2h. Cette approche est justifiée par le besoin de lire le marqueur à des distances différentes malgré une précision de l'information décroissante avec l'éloignement [71]. Les hautes fréquences encodent ainsi les informations les plus précises, alors que les basses fréquences toujours lisibles malgré une résolution de l'image plus faible encodent l'information importante.

1.6 Conclusion et présentation d'un marqueur spécifique, le bleam

Dans ce chapitre nous avons décrit le fonctionnement général des systèmes de marqueurs fiduciaires, ainsi que les méthodes et les algorithmes nécessaires à leurs mises en place.

Les différents marqueurs se distinguent par la façon de les détecter dans l'image. Nous avons vu qu'ils peuvent utiliser un seuillage pour détecter leur emplacement dans l'image ou une convolution par un filtre de détection de contours de type Sobel ou Canny [9]. Les ellipses, projections du contour des marqueurs circulaires, sont des primitives qui peuvent être estimées de façon fiable même si tous leurs points ne sont pas visibles sur une vue.

Les marqueurs se distinguent également par la méthode utilisée pour estimer l'homographie. Les marqueurs carrés utilisent les correspondances estimées dans l'image pour les 4 sommets du carré. Cette méthode permet de retrouver entièrement l'homographie

mais la qualité de cette estimation dépend de la précision de la détection des coins du quadrilatère dans l'image. À l'inverse, l'utilisation du contour des marqueurs circulaires n'est pas suffisant pour calculer l'homographie, dans son intégralité, mais l'ellipse qui correspond à l'image de son contour qui peut être estimée de façon robuste fournit de solides contraintes sur cette homographie. Enfin les marqueurs se distinguent surtout par la façon dont ils encodent leur identifiant. Chaque système de marqueur utilise généralement un système unique, on retrouve souvent cependant un encodage binaire avec un code cyclique de redondance.

Dans la suite de ce manuscrit, nous allons particulièrement nous intéresser aux marqueurs circulaires, car ils présentent une robustesse :

- aux bruits présents dans l'image [1] ;
- aux occultations (de leur image) [1] ;
- au flou de mouvement [7].

En particulier nous allons travailler avec des marqueurs de type *bleam* dont une illustration est présentée sur la FIGURE 1.6. Ce marqueur est constitué d'un anneau externe

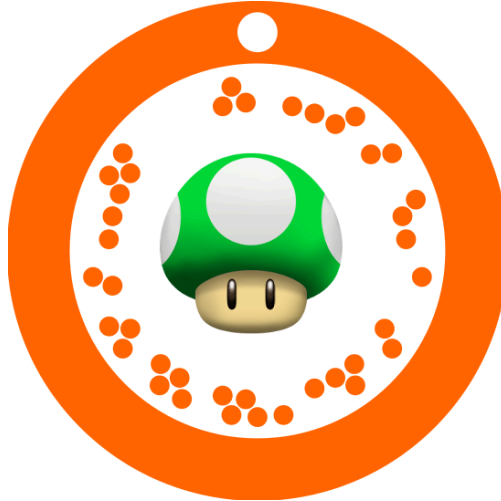


FIGURE 1.6 – Le *bleam*.

noir ou coloré constitué de deux cercles. Le rapport entre les rayons du cercle externe et du cercle interne est de 0.75. Un détrompeur circulaire est placé sur l'anneau pour identifier la rotation du marqueur autour de son centre. Deux anneaux de points encodent un code binaire de 80 bits à l'intérieur du marqueur. Enfin, généralement un logo ou une image peut être placé au centre du marqueur. Le marqueur dispose ainsi de suffisamment de contraintes sur son apparence afin d'être rectifié et identifié.

Dans le prochain chapitre, nous allons présenter et étudier les méthodes existantes de calcul de l'homographie à partir de primitives géométriques et en particulier de cercles. Ce chapitre nous permettra de mettre en avant les méthodes pertinentes pour le calcul de l'homographie à partir de marqueurs circulaires.

Chapitre 2

Calcul de l'homographie pour les marqueurs circulaires

Sommaire

2.1	Calcul de l'homographie de rectification par sa décomposition	22
2.2	Méthodes d'estimation de l'homographie	23
2.2.1	À partir de points et de droites	24
2.2.2	À partir de plusieurs cercles ou ellipses coplanaires	25
2.2.3	À partir de cercles concentriques	30
2.2.4	À partir d'un cercle, dans le cas calibré	35
2.3	Application à un cas spécifique : le contour d'un bleam . .	36
2.3.1	Présentation des méthodes comparées	37
2.3.2	Résultats des tests sur la rectification	41
2.4	Conclusion sur les méthodes et perspectives	42

2.1 Calcul de l'homographie de rectification par sa décomposition

Comme nous l'avons vu dans le chapitre 1, l'intérêt principal des marqueurs plans est qu'à partir de leurs motifs, nous disposons de suffisamment de primitives pour estimer l'homographie. Dans ce chapitre, nous parlons d'homographie pour qualifier la perspective qui est la restriction au plan de support, noté \mathcal{P} , du marqueur, de la projection par une caméra. Nous nous plaçons dans le cadre d'une projection quasi-affine, c'est-à-dire que tous les points ou toutes les primitives considérées se situent devant la caméra. Dans ce contexte, les cercles ont toujours pour image des ellipses.

Nous rappelons que la perspective, comme toute homographie, peut être décomposée en trois transformations particulières [30, p.43] :

$$H = H_P H_A H_S = \begin{bmatrix} I_2 & \mathbf{0} \\ \mathbf{v}^T & 1 \end{bmatrix} \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.1)$$

Les trois transformations sont respectivement une transformation perspective, une transformation affine et une similitude. À noter que \mathbf{v} , K , s , R , et \mathbf{t} ne sont pas des paramètres indépendants et doivent se distinguer des paramètres de la caméra.

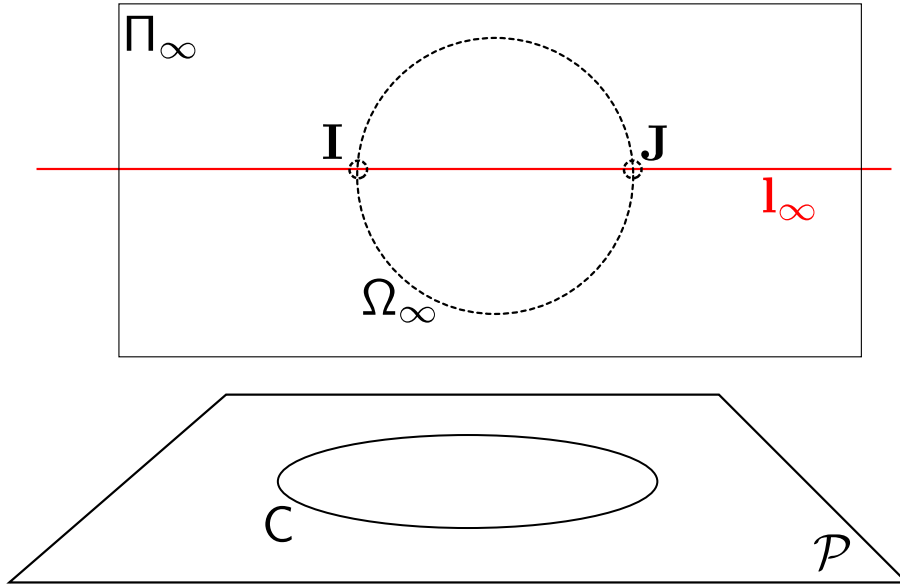


FIGURE 2.1 – Illustration des éléments caractéristiques de l'espace projectif.

Pour la suite de ce chapitre, nous invitons le lecteur qui ne serait pas familier avec les propriétés élémentaires de la géométrie projective à lire l'annexe A. Afin de clarifier la lecture nous donnons une illustration des objets géométriques caractéristiques de l'espace projectif sur la FIGURE 2.1.

La rectification affine consiste à trouver une transformation qui transforme la ligne de fuite en la droite à l'infini du plan projectif sous sa forme canonique, notée \mathbf{l}_∞ sur

la FIGURE 2.1. Cette rectification permet donc de retrouver le parallélisme des droites du plan de support après leurs projections sur la photographie ; les droites parallèles s'intersectant sur cette ligne de fuite. Du point de vue de la décomposition présentée dans l'équation 2.1, ceci consiste à retrouver la première transformation H_P où \mathbf{v} est la ligne de fuite.

La rectification métrique est une transformation homographique particulière qui permet de remettre sous la forme canonique d'autres primitives particulières du plan de support. Nous admettons que les trois définitions suivantes sont équivalentes. Étant donné un plan 3D et un plan image en relation de perspective, une rectification métrique :

- transforme l'image d'un carré en carré ;
- transforme l'image de deux cercles en deux cercles ;
- transforme les coordonnées de l'image des points cycliques en leurs coordonnées canoniques.

On rappelle que les points cycliques sont deux points complexes conjugués notés \mathbf{I} et \mathbf{J} sur la FIGURE 2.1. Ces points peuvent être définis comme l'intersection commune de tous les cercles du plan, plus d'explications à ce sujet sont données en annexe A.8.

Du point de vue de la décomposition présentée dans l'équation 2.1, rectifier affinement consiste à identifier les deux premières transformations H_P et H_A de notre perspective. Ainsi une rectification métrique transforme également la ligne de fuite en la droite à l'infini de coordonnées canoniques. Le calcul de cette rectification est un premier pas vers le calcul de l'homographie.

Dans ce chapitre, nous nous intéressons au calcul de l'homographie en utilisant différentes primitives (points, segments, cercles). Dans un premier temps, nous présentons les méthodes de l'état de l'art sur les primitives circulaires à l'instar du contour du marqueur que nous avons présenté. Celui-ci présente deux cercles concentriques. Nous nous attacherons donc particulièrement aux méthodes permettant le calcul de l'homographie avec un ou deux cercles concentriques. Finalement nous retiendrons deux de ces méthodes d'estimation que nous évaluerons après les avoir adaptées aux spécificités de notre problème.

2.2 Méthodes d'estimation de l'homographie

Nous rappelons qu'une homographie permet de caractériser la projection du plan de support du marqueur par une caméra et peut être entièrement paramétrée par une matrice d'homographie d'ordre 3, que nous noterons H . De plus, il est possible de calculer ses coefficients en appliquant des contraintes entre les primitives géométriques attendues (carrés, cercles) et leurs observations sur la vue. En exprimant les contraintes apportées par les correspondances entre les primitives et leurs images, le problème peut se formuler ainsi :

$$\mathcal{H}(P_i) = P'_i \text{ avec } P'_i \text{ l'image de la primitive } P_i. \quad (2.2)$$

La matrice correspondant à l'homographie possède neuf coefficients. Comme il s'agit d'une représentation homogène, elle possède huit degrés de liberté. Selon le type de primitives utilisées, le nombre d'équations linéairement indépendantes obtenues varie ce qui change le nombre de degrés de liberté de l'homographie. Une correspondance de points ou de droites (équations similaires à cause de leur dualité dans le plan projectif) fournit deux équations linéaires indépendantes. Une correspondance de coniques fournit cinq équations linéairement indépendantes sur l'homographie. De façon générale, une

correspondance de courbes polynomiales de degré n réduit au plus de $\frac{1}{2}n(n+3)$ le nombre de degrés de liberté de l'homographie [1].

À noter que lorsque la photographie du marqueur est connue, il est également possible de calculer directement l'homographie en minimisant la différence photométrique avec l'image rectifiée du marqueur observé. Cette classe de méthodes n'est cependant pas le sujet de ce chapitre. En effet, dans notre cas, les primitives présentes sur le marqueur utilisé simplifient le calcul de l'homographie et permettent donc de calculer l'homographie sans recours à une comparaison photométrique avec l'image du marqueur.

Par la suite, nous distinguons les méthodes en fonction des différents types de primitives utilisées. Comme expliqué en annexe A.7.2, pour calculer l'homographie, nous pouvons, dans un premier temps, simplifier le problème, en calculant la rectification métrique ou la rectification affine (calcul des invariants).

Certaines méthodes présentées se concentrent ainsi sur le calcul des invariants, ce qui nécessite moins de correspondances. Toutefois, l'homographie ne sera estimée qu'à une similitude ou à une transformation affine près. Nous étudierons dans un premier temps les méthodes utilisant des correspondances de points puis nous verrons comment calculer l'homographie à partir de correspondances de cercles coplanaires. Nous verrons ensuite le cas particulier des cercles concentriques puis finalement nous aborderons le cas où les paramètres intrinsèques sont disponibles. En effet, lorsque la caméra est calibrée il est possible d'obtenir l'homographie à partir d'une seule correspondance de cercles à une rotation près.

2.2.1 À partir de points et de droites

Lorsque des correspondances de droites et de points entre plusieurs vues sont connues, il est possible de lier linéairement les coefficients de l'homographie avec ces données. Quatre correspondances de points ou de droites sont suffisantes pour obtenir une solution exacte du problème. À partir de $n \geq 4$ correspondances de points (ou de droites), le problème est nommé *DLT* (de l'anglais *Direct Linear Transform*) [30, p. 90-91], il peut être résolu en utilisant une stratégie de type moindre carré. En coordonnées homogènes, la correspondance de deux droites ou de deux points permet de contraindre l'homographie de la même façon.

$$\mathbf{H}\mathbf{x} \sim \mathbf{x}' \quad (2.3)$$

$$\mathbf{H}^{-\top}\mathbf{l} \sim \mathbf{l}' \quad (2.4)$$

avec \mathbf{x}, \mathbf{x}' les vecteurs homogènes d'un point et son image et \mathbf{l}, \mathbf{l}' les vecteurs homogènes d'une droite et son image. Nous pouvons ainsi écrire :

$$[\mathbf{x}']_{\wedge} \mathbf{H}\mathbf{x} = 0 \quad (2.5)$$

Cela nous donne ainsi un ensemble d'équations qui dépend linéairement des coefficients de l'homographie :

$$\begin{bmatrix} \mathbf{0}^{\top} & -x'_3\mathbf{x}^{\top} & x'_2\mathbf{x}^{\top} \\ x'_3\mathbf{x}^{\top} & \mathbf{0}^{\top} & -x'_1\mathbf{x}^{\top} \\ -x'_2\mathbf{x}^{\top} & x'_1\mathbf{x}^{\top} & \mathbf{0}^{\top} \end{bmatrix} \begin{bmatrix} \mathbf{H}^1 \\ \mathbf{H}^2 \\ \mathbf{H}^3 \end{bmatrix} = \mathbf{0} \quad (2.6)$$

De façon similaire, nous pouvons généraliser ces équations au cas des droites, il suffit d'interchanger les lignes pour prendre en considération la transposée de \mathbf{H} . Les trois équations ne sont pas indépendantes, seulement deux le sont, ce qui justifie le fait qu'il

est nécessaire d'avoir quatre correspondances (il est possible de compléter le problème, soit en fixant un coefficient de l'homographie, soit en ajoutant une contrainte sur la norme du vecteur).

2.2.2 À partir de plusieurs cercles ou ellipses coplanaires

Les données des cercles et leurs images par projection perspective fournissent également des contraintes sur l'homographie entre le plan de support et le plan image. Ceci se généralise pour toute conique du plan projectif (cf. annexe A.5). Avec deux correspondances de coniques du plan projectif complexifié, c'est-à-dire en remplaçant le corps des réels par le corps des complexes, nous savons que ces deux coniques possèdent quatre points d'intersection. En effet l'appartenance aux deux coniques est équivalente à résoudre le système d'équation polynomiale à 2 polynômes homogènes de degré 2 (une pour chaque conique) sur les coordonnées homogènes du point. Le théorème de Bézout stipule en effet que si les deux polynômes homogènes n'ont pas de composantes communes, les courbes correspondantes se coupent en 2×2 points. Nous verrons que dans cette section ceci est vrai pour des cercles ou des ellipses quelconques. Cependant pour des cercles ou des ellipses ayant des configurations particulières, les équations peuvent avoir des composantes communes (cas des cercles concentriques) et ainsi ne pas entièrement contraindre une homographie. Mais dans le cas général, il est possible d'estimer une homographie. On rappelle qu'il faut quatre correspondances de points pour estimer l'homographie. Pour rendre le problème linéaire ou simplifier sa résolution, certains auteurs proposent également d'utiliser plus de deux correspondances de cercles ou d'ellipses. Le paragraphe suivant nous permet d'exposer certaines de ces solutions.

Calcul direct de l'homographie

Sugimoto *et al.* [75] proposent une méthode linéaire pour estimer la matrice d'homographie à partir de sept correspondances de coniques. En partant du constat qu'une conique du plan projectif \mathbb{P}^2 peut être représentée par un point de \mathbb{P}^5 , il s'ensuit que la transformation homographique d'une conique de \mathbb{P}^2 peut être représentée par une transformation linéaire dans \mathbb{P}^5 où la matrice associée est d'ordre 6. Sept correspondances permettent donc de calculer cette transformation linéaire pour retrouver cette transformation de \mathbb{P}^5 dans \mathbb{P}^5 . Chaque correspondance de point de \mathbb{P}^5 permet en effet de déterminer cinq degrés de liberté de la matrice d'ordre 6. Les auteurs montrent ensuite qu'il est possible d'extraire des équations linéaires sur les coefficients de l'homographie plane en utilisant les coefficients de la matrice d'ordre 6 associés à la transformation sur \mathbb{P}^5 . Malheureusement, dans le cas des marqueurs, cette méthode nécessite cependant trop de correspondances pour être utilisée en pratique et présente uniquement un intérêt théorique. Elle peut cependant être utilisée à des fins de calibrage de caméra.

Mudigonda *et al.* [54] montrent qu'il est possible de calculer l'homographie à partir de deux coniques coplanaires en correspondance dans deux vues. Ils proposent de traiter le problème en calculant les deux homographies de rectification métrique qui transforment chacune des paires de coniques en paire de cercles. L'homographie qui permet de passer d'une vue à une autre peut ensuite être calculée en composant les deux perspectivités.

Dans le cas particulier de deux cercles du plan projectif complexifié, leur intersection donne deux paires de points dont l'une est la paire des points complexes conjugués qui correspond aux points cycliques. En effet, comme expliqué en introduction du chapitre, les points cycliques sont contenus dans tous les cercles du plan. L'intersection des deux coniques (images des cercles) permet donc d'obtenir les images des points cycliques. De

plus, si les coordonnées des cercles sont connues, l'autre paire de points d'intersection permet d'obtenir suffisamment de correspondances pour calculer la similitude plane, ce qui permet de déterminer entièrement l'homographie de chacune des perspectives.

Cette méthode nécessite que les coniques soient de rang plein. En effet, nous rappelons que la projection conserve le rang des coniques. Il est donc nécessaire de manipuler des coniques de même rang qu'un cercle, c'est-à-dire correspondant à des ellipses ou des hyperboles. Comme nous obtenons deux équations quadratiques à deux inconnues, la méthode nécessite également de résoudre une équation quartique pour calculer les points d'intersection, puisqu'on obtient un ensemble de deux équations quadratiques à deux inconnues. Les quatre solutions obtenues apportent une ambiguïté, car la paire de points cycliques est indistinguable de l'autre paire de points complexes qui peut également être solution.

Kannala *et al.* proposent également de résoudre le problème avec $N > 3$ coniques en correspondance. Ils utilisent pour cela une résolution par les moindres carrés linéaires [39]. Plus précisément, pour chaque couple de coniques (C_1, C_2) et son image (E_1, E_2) , il est possible d'obtenir ces deux équations :

$$C_1 \sim H^\top E_1 H \quad (2.7)$$

$$C_2 \sim H^\top E_2 H \quad (2.8)$$

À partir de ces équations, nous obtenons :

$$E_1^{-1} E_2 H - H C_1^{-1} C_2 = 0 \quad (2.9)$$

Cette équation donne cinq contraintes linéaires sur la matrice de l'homographie H . Ainsi, à partir de trois coniques, nous obtenons quinze contraintes linéaires qui permettent d'apporter une solution linéaire. Les auteurs introduisent également une solution pour deux coniques où ils proposent de factoriser les matrices symétriques des coniques avec les matrices complexes F_1 et F'_1 :

$$C_1 \sim F_1^\top F_1$$

$$E_1 \sim F_1'^\top F'_1$$

Ainsi, en utilisant les équations de projections (2.7), ils montrent que les deux coniques C_1 et C_2 ont pour image E_1 et E_2 par la même homographie si et seulement si :

$$R^\top R = I_3 \text{ et } R^\top A R \sim B \quad (2.10)$$

avec :

$$A = F_1'^{-1} E_2 F_1'^{-\top} \text{ et } B = F_1^{-1} C_2 F_1^{-\top}$$

Autrement dit, si A et B sont similaires. En réalisant la décomposition en vecteurs propres de A et B , il est possible de trouver une solution pour R . Enfin, les auteurs montrent que H peut alors être calculée avec :

$$H = F_1'^{-\top} R F_1^\top \quad (2.11)$$

À noter que ce problème revient à résoudre un problème de calcul de valeurs propres et vecteurs propres généralisés et la solution peut ainsi être exprimée en calculant les vecteurs propres généralisés de A et B . Dans le cas de cercles concentriques, A et B

partagent une valeur propre de multiplicité double. Ainsi, la solution pour R n'est pas unique. Ce qui se comprend géométriquement, l'image de deux cercles concentriques reste inchangée par une rotation du plan de support centrée sur les deux cercles. La méthode peut cependant être intéressante pour calculer l'homographie dans le cadre de coniques quelconques (de rang plein et sans valeurs propres communes de multiplicité double).

Une méthode plus récente proposée par Huang *et al.* [36] fournit une estimation de l'homographie à partir de deux ellipses coplanaires non incluses l'une dans l'autre et leurs images. Les auteurs utilisent la notion de triangle auto-polaire à une ellipse. Plus précisément, cette notion désigne un triangle dont tous les sommets sont en relation polaire par rapport au cercle avec le côté qui leur est opposé, cette relation est expliquée en annexe A.6. Les auteurs montrent que pour deux ellipses disjointes, il n'existe qu'un seul triangle auto-polaire commun. Celui-ci contient deux sommets dans chacune des ellipses et un sommet à l'extérieur des deux. Les droites associées au segment de ce triangle peuvent être calculées à partir des vecteurs propres de $C_2^{-1}C_1$, où C_1 et C_2 sont les matrices des deux ellipses. En calculant également l'image de ce triangle à partir des images de ces coniques, les auteurs extraient ainsi trois correspondances de droites. Ils montrent également que ce triangle intersecte les ellipses toujours en deux points. Ainsi, ces deux points fournissent une quatrième correspondance de droites nécessaire au calcul de l'homographie.

Calcul préliminaire de la rectification métrique

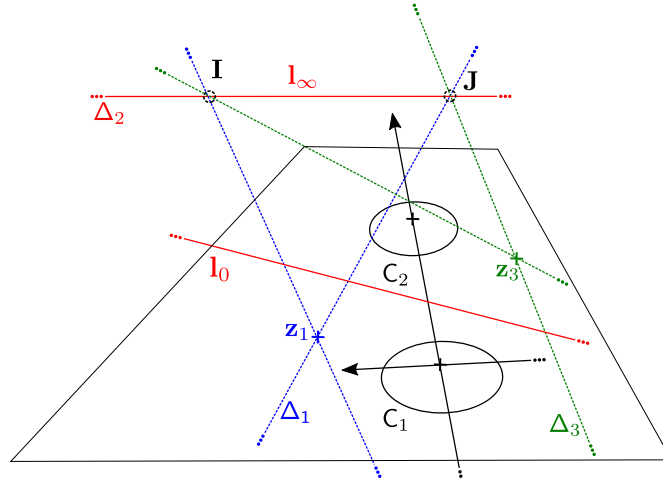


FIGURE 2.2 – Représentation du plan projectif contenant deux cercles avec les trois coniques dégénérées du faisceau. La droite à l'infini passe par les deux points cycliques, points d'intersection des deux cercles.

Dans le cas général de l'image de deux cercles coplanaires (ou parallèles) disjoints, il est possible de retrouver l'homographie en calculant, dans un premier temps, la rectification métrique entre les images. Dans cette section, nous présentons donc les méthodes qui traitent ce problème. Considérons deux cercles coplanaires d'un plan euclidien, de matrices C_1 et C_2 , sans perte de généralité, nous pouvons choisir un repère tel que le centre du premier cercle est à l'origine et l'axe Ox passe par le centre du second cercle.

Dans ce repère, les matrices des cercles peuvent s'écrire ainsi :

$$\mathbf{C}_1 \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r_1^2 \end{bmatrix} \text{ et } \mathbf{C}_2 \sim \begin{bmatrix} 1 & 0 & -d \\ 0 & 1 & 0 \\ -d & 0 & d^2 - r_2^2 \end{bmatrix} \quad (2.12)$$

d est la distance entre les deux centres et r_1, r_2 sont les rayons respectifs des cercles. Ainsi, les points cycliques sous forme canonique ($\mathbf{I}, \mathbf{J} = [1, \pm i, 0]$) passent par les deux cercles.

Nous pouvons définir un faisceau de coniques à partir de la paire de cercles ($\mathbf{C}_1, \mathbf{C}_2$), par l'ensemble des coniques ayant pour matrice $\mathbf{C}_1 - \lambda \mathbf{C}_2$ avec $\lambda \in \mathcal{R}$. À noter que nous nous permettons d'avoir $\lambda = \infty$, ce qui permet d'inclure \mathbf{C}_2 dans le faisceau.

Gurdjos *et al.* montrent que le faisceau de coniques contient trois coniques dégénérées de rang 2. Toute conique du faisceau contient les points communs aux deux cercles. Toute conique du faisceau contient donc les points cycliques y compris les coniques dégénérées. Deux des coniques dégénérées de matrices Δ_1 et Δ_3 sont des coniques composées d'un seul point réel (conique de signature (2,0) dans le TABLEAU A.1 fourni en annexe) respectivement \mathbf{z}_1 et \mathbf{z}_3 . Ces coniques sont composées de deux droites complexes passant chacune par le point réel et un des points cycliques. Nous pouvons noter que ces droites complexes sont isotropes et permettent d'obtenir des contraintes sur les points cycliques de vecteurs, \mathbf{I}, \mathbf{J} . La troisième conique dégénérée notée Δ_2 sur la figure est composée de deux droites réelles dont l'une est la droite à l'infini, \mathbf{l}_∞ , elle a pour signature (1,1). Une illustration de la configuration est donnée dans la FIGURE 2.2.

Comme expliqué dans l'annexe A.5, quand nous étudions les projections dans le plan image, les signatures des coniques sont conservées. Ainsi, si \mathbf{E}_1 et \mathbf{E}_2 sont respectivement l'image de \mathbf{C}_1 et \mathbf{C}_2 , la conique dégénérée de rang 2 du faisceau $\mathbf{E}_1 - \lambda \mathbf{E}_2$ ayant pour signature (1,1) est la conique contenant l'image de la droite à l'infini, la ligne de fuite.

Cependant, une difficulté persiste car il est nécessaire de distinguer les différentes coniques dégénérées obtenues en classant les valeurs propres ou en distinguant la signature des coniques dégénérées. La rectification affine peut être retrouvée à une ambiguïté près, il faudra en effet prendre en compte les deux droites de l'image de la conique Δ_2 qui, sans information supplémentaire, peuvent être considérées comme étant la ligne de fuite. La rectification métrique est ensuite réalisable en utilisant les deux autres coniques dégénérées qui imposent une contrainte sur l'image des points cycliques.

Chen *et al.* [10, 34] expliquent qu'à partir d'un modèle simplifié d'une caméra paramétrée uniquement par la distance focale, il est possible de calculer simultanément les paramètres intrinsèques et la pose de la caméra relative au plan observé. Pour cela, il propose d'utiliser également deux coniques coplanaires. L'hypothèse est que seul le paramètre de la focale est inconnu, le reste des paramètres intrinsèques de la caméra est supposé connu. Ainsi, la matrice des paramètres intrinsèques peut s'écrire :

$$\mathbf{K} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \quad (2.13)$$

Les auteurs expriment dans deux repères différents le cône passant par le centre optique et un des cercles observés, le repère caméra et un repère orienté par le plan de support du cercle. La FIGURE 2.3 illustre ce cône et les deux repères associés à chacun des plans considérés. Le cône dans le repère caméra peut être paramétré par l'ellipse observée située sur le plan image, \mathcal{I} , et le centre optique est fixé à l'origine du repère.

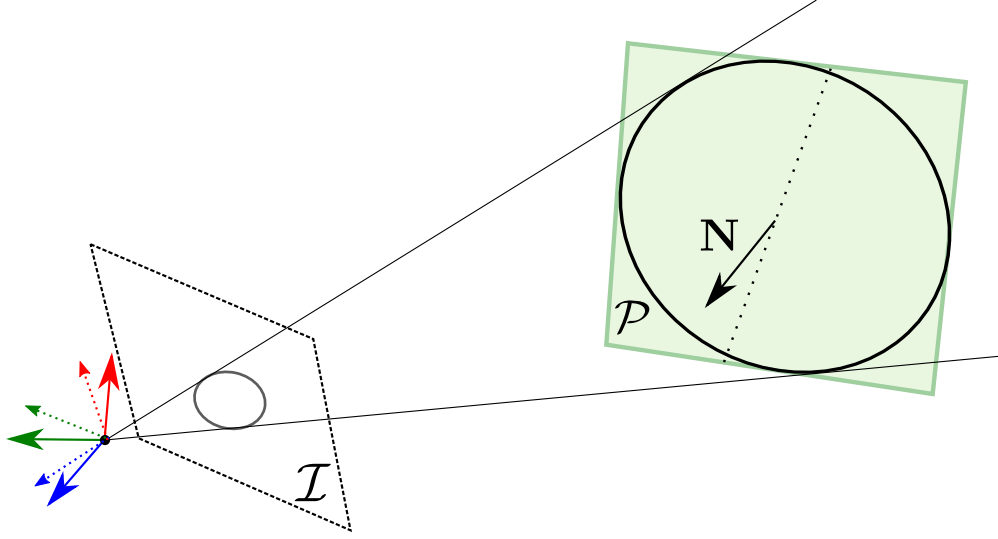


FIGURE 2.3 – Le cône formé du centre optique et du cercle du plan de support, \mathcal{P} , intersecte également la projection elliptique du cercle sur le plan image, \mathcal{I} .

Nous notons E_1 , la matrice de l'ellipse observée dans le plan image dont le plan a pour équation $z = -f$ où f est la focale de la caméra, supposée connue (ou égale à 1 pour initialiser une solution). Si nous supposons la matrice de projection P dans le repère caméra, celle-ci a pour équation :

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix} [I \quad \mathbf{0}] \quad (2.14)$$

avec I la matrice identité d'ordre 3 et $\mathbf{0}$ un vecteur nul de dimension 3.

Ainsi, dans le repère caméra le cône passant par l'origine a pour matrice :

$$Q_c \sim P^T E_1 P \quad (2.15)$$

Cette matrice s'écrit :

$$Q_c \sim \begin{bmatrix} \bar{Q}_c & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \text{ avec } \bar{Q}_c \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix}^T E_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \quad (2.16)$$

Considérons maintenant le repère avec pour origine, toujours, le centre de la caméra mais orienté en z par la normale du plan de support du cercle. Celui-ci a ainsi pour équation $z = z_0$. Dans ce repère, le cône passant par le cercle de centre $[x_0 \ y_0 \ z_0]^T$ et de rayon r a pour matrice :

$$Q \sim \begin{bmatrix} \bar{Q} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \text{ avec } \bar{Q} \sim \begin{bmatrix} 1 & 0 & -\frac{x_0}{z_0} \\ 0 & 1 & -\frac{y_0}{z_0} \\ -\frac{x_0}{z_0} & -\frac{y_0}{z_0} & \frac{x_0^2 + y_0^2 - r^2}{z_0^2} \end{bmatrix} \quad (2.17)$$

Les deux matrices associées au même cône ne diffèrent que par le changement de repère. Ces deux repères étant seulement liés par une rotation de matrice R , passant du

repère orienté selon le plan de support vers le repère caméra, nous avons :

$$\mathbf{Q} \sim \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{Q}_c \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}^\top & 1 \end{bmatrix}^\top \text{ ainsi } \bar{\mathbf{Q}} \sim \mathbf{R}^\top \bar{\mathbf{Q}}_c \mathbf{R} \quad (2.18)$$

Cette équation peut être résolue en utilisant une décomposition en valeurs singulières de $\bar{\mathbf{Q}}_c = \mathbf{V}^\top \Sigma \mathbf{V}$, les auteurs obtiennent ainsi des contraintes sur la rotation qui leur permettent ainsi d'obtenir la normale du plan de support [10] :

$$\mathbf{N} \sim \mathbf{V} \begin{bmatrix} \pm \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} \\ 0 \\ \pm \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \end{bmatrix} \quad (2.19)$$

où $(\lambda_1, \lambda_2, \lambda_3)$ sont les valeurs propres triées par ordre décroissant de $\bar{\mathbf{Q}}_c$ et les signes \pm symbolisent l'ambiguïté qui existe dans les solutions. En utilisant ce procédé sur deux cercles coplanaires, l'ambiguïté peut être fixée. La focale peut être déterminée en paramétrant les deux normales obtenues \mathbf{N}_1 et \mathbf{N}_2 par la distance focale et en minimisant :

$$||\mathbf{N}_1(f)^\top \mathbf{N}_2(f) - 1||$$

La méthode ne précise pas l'algorithme d'optimisation utilisé et nous pouvons noter que l'expression de \mathbf{N} est fortement non linéaire en f , il est donc nécessaire d'utiliser un algorithme itératif pour atteindre un optimum local. De plus, aucune preuve de convergence globale n'est disponible, il est important de bien initialiser la distance focale f à une valeur proche des données réelles.

La méthode proposée par Chen *et al.* a l'avantage de fonctionner sans condition sur la position relative des cercles puisque la résolution se fait indépendamment pour chaque cercle. Elle fonctionne donc aussi pour des cercles concentriques et ne nécessite que deux correspondances de cercles. Cependant, la méthode repose sur des *a priori* très forts sur les paramètres intrinsèques de la caméra pour réaliser l'optimisation nécessaire.

2.2.3 À partir de cercles concentriques

Le marqueur que nous avons considéré dans ces travaux, présenté à la fin du chapitre 1, est formé de deux cercles concentriques. Les solutions vues précédemment ne sont pas adaptées à ce type de marqueur. Ainsi, dans la suite de cette section, nous détaillons les méthodes qui traitent ce cas spécifiquement. La première étape à considérer est la rectification affine, partie la plus délicate du problème. Une fois cette rectification obtenue, une méthode simple permet de calculer une solution pour l'homographie entre le plan de support et le plan image. À noter qu'il n'existe pas une solution unique, une rotation des cercles dans le plan de support autour de l'axe passant par leur centre commun ne change pas leurs images dans la photographie. La FIGURE 2.4 illustre une configuration géométrique possible de la projection de deux cercles concentriques. Nous pouvons y observer des objets particuliers du plan qui vont nous être utiles dans le reste de cette section. Il s'agit, par exemple, de la droite à l'infini, notée \mathbf{l}_∞ qui donne la structure affine du plan et des points cycliques qui permettent de définir la structure euclidienne. Le point de vecteur de coordonnées \mathbf{c} est le centre des deux cercles.

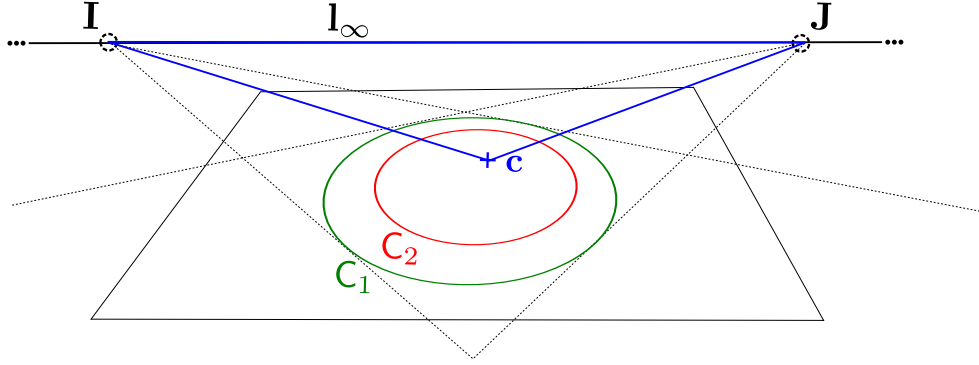


FIGURE 2.4 – Projection de deux cercles concentriques. Leur centre commun, le point de vecteur \mathbf{c} est en relation polaire avec la droite à l'infini, de vecteur \mathbf{l}_∞ , par rapport à chacun des deux cercles.

Calcul préliminaire de la structure affine

Dans un premier temps nous présentons le raisonnement suivi par Kim *et al.* [40]. Considérons deux cercles concentriques du plan, C_1 et C_2 , sans perte de généralité, nous fixons le centre du repère sur le centre des cercles, leurs matrices respectives peuvent ainsi s'écrire sous la forme :

$$C_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r_i^2 \end{bmatrix} \text{ avec } i \in \{1, 2\} \text{ et } r_i, \text{ rayon du cercle.} \quad (2.20)$$

Gurdjos *et al.* [26] proposent d'étudier le problème dans l'espace dual. Ainsi, la conique de points C_i est remplacée par sa duale, de matrice C_i^{-1} (composée de ses droites tangentes).

Si nous effectuons le calcul des coniques duales aux cercles concentriques, nous obtenons deux nouvelles coniques :

$$C_i^{-1} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{r_i^2} \end{bmatrix} \quad (2.21)$$

Étudions maintenant le faisceau de cercles $C_1^{-1} - \lambda C_2^{-1}$ avec $\lambda \in \bar{\mathcal{R}}$. Nous pouvons vérifier que ce faisceau est composé de cercles réels, de cercles virtuels¹ et de trois cercles dégénérés. Concernant les coniques dégénérées, il en existe deux identiques de rang 1 et une autre de rang 2 [40].

Lorsque $\lambda = 1$, nous pouvons vérifier que nous avons :

$$\tilde{\Delta}_1 \sim \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \mathbf{c} \mathbf{c}^\top \quad (2.22)$$

1. Un cercle virtuel est un cercle dont la matrice est réelle, mais qui ne contient pas de points réels.

Cette conique dégénérée correspond à un unique point qui est le centre des cercles, noté \mathbf{c} . Ainsi, cette conique décrit toutes les droites passant par ce point unique. Kim *et al.* la nomment la *conique duale au centre du cercle* [40]. Cette conique a donc une grande importance pour calculer une rectification affine de l'image. En effet, lors de sa projection, elle contiendra l'image du centre des cercles qui est en relation polaire avec la ligne de fuite du plan (voir la définition A.6.2 dans l'annexe A.6).

Lorsque $\lambda_2 = \frac{r_2^2}{r_1^2}$, nous avons la seconde conique qui est la conique duale aux points cycliques et elle est composée des deux points cycliques. Cette conique particulière, à savoir la conique duale des points cycliques, notée IJ [30, p. 53], contient toutes les droites incidentes aux points cycliques. La connaissance de cette conique donne directement les points cycliques par décomposition en valeurs singulières de sa matrice.

Nous pouvons ainsi écrire :

$$\tilde{\Delta}_2 \sim \mathbf{C}_\infty^* \sim \mathbf{IJ}^\top + \mathbf{JI}^\top \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (2.23)$$

Ainsi, si \mathbf{E}_1 et \mathbf{E}_2 sont respectivement l'image de \mathbf{C}_1 et \mathbf{C}_2 , la conique dégénérée de rang 2 du faisceau $\mathbf{E}_1^{-1} - \lambda_2 \mathbf{E}_2^{-1}$ avec $\lambda \in \bar{\mathcal{R}}$ est l'image de la conique duale aux points cycliques. Nous rappelons que les coniques dégénérées avec lesquelles nous travaillons sont dans l'espace dual et il s'agit donc de coniques de droites. Ainsi, elles peuvent être définies par l'ensemble des droites tangentes. La FIGURE 2.5 donne une illustration des coniques $\tilde{\Delta}_1$ et $\tilde{\Delta}_2$. La conique dégénérée de matrice $\tilde{\Delta}_1$ correspond à l'ensemble des droites du plan qui passent par le centre des cercles. La conique dégénérée de matrice $\tilde{\Delta}_2$ correspond à l'ensemble des droites qui passent par au moins un des points cycliques. On peut noter que la droite à l'infini, de vecteur \mathbf{l}_∞ est incluse dans cette conique.

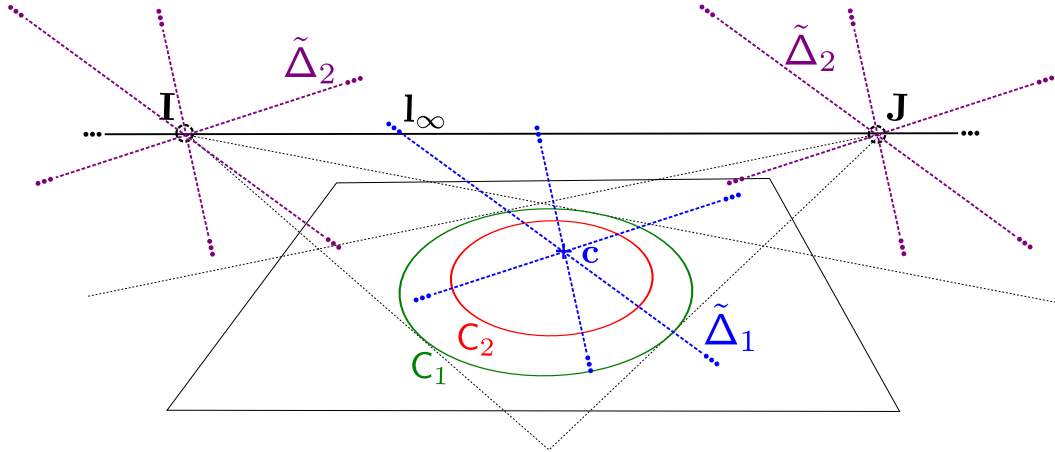


FIGURE 2.5 – Les deux coniques dégénérées du faisceau $\{\mathbf{C}_1^{-1} - \lambda \mathbf{C}_2^{-1}\}$. La conique dégénérée de matrice $\tilde{\Delta}_1$ correspond à l'ensemble des droites passant par le centre du cercle. La seconde conique, de matrice $\tilde{\Delta}_2$ correspond à l'ensemble des droites passant par au moins un des points cycliques.

L'algorithme de rectification proposé par Kim *et al.* consiste à calculer la rectification affine à partir de l'image de la conique de rang 1, $\tilde{\Delta}_1$ que nous noterons $\tilde{\mathbf{D}}_1$. Cette conique a la particularité de n'être composée que d'un point réel double, l'image du centre du

cercle :

$$\tilde{D}_1 \sim \mathbf{x}_c^\top \mathbf{x}_c. \quad (2.24)$$

Il a été montré que les valeurs propres généralisées de (E_1^{-1}, E_2^{-1}) , à condition de normaliser les matrices de E_1^{-1} et E_2^{-1} de sorte que leur déterminant valent 1, sont les mêmes que celles de (C_1^{-1}, C_2^{-1}) [27]. Nous pouvons ainsi retrouver l'image de la conique dégénérée \tilde{D}_1 en triant les valeurs propres généralisées du couple (E_1^{-1}, E_2^{-1}) , $\mu_1 \leq \mu_2 \leq \mu_3$. Celle associée à l'image de la conique dégénérée de rang 2, \tilde{D}_2 , notée \tilde{D}_2 , a pour multiplicité 2 ainsi $E_1^{-1} - \mu_2 E_2^{-1}$ est nécessairement l'image de cette conique. Ceci permet donc d'obtenir la conique \tilde{D}_1 , en sélectionnant la valeur propre différente de μ_2 . En calculant le seul vecteur propre de valeur propre non nulle de la matrice \tilde{D}_1 , nous obtenons l'image du centre \mathbf{x}_c . Enfin, grâce à la relation polaire, nous pouvons obtenir la ligne de fuite. Nous rappelons que $\mathbf{v}_\infty = E_1 \mathbf{x}_c$. L'algorithme 1 présente ce calcul. La condition ligne 5 de l'algorithme permet de trouver la valeur propre double. Comme il existe deux valeurs propres proches mais distinctes à cause des erreurs de mesures ou de calcul flottant, nous considérons la moyenne des deux valeurs propres les plus proches.

Algorithm 1 `rectifAffine_EE` : Calcul de la ligne de fuite à partir de deux cercles concentriques. La fonction *geigen* correspond aux valeurs propres généralisées.

Entrée: E_1, E_2

Sortie: \mathbf{v}_∞

1: $E_1 \leftarrow \text{signe}(\det(E_1))E_1 / \sqrt[3]{ \det(E_1) }$	\triangleright Normalisation des coniques
2: $E_2 \leftarrow \text{signe}(\det(E_2))E_2 / \sqrt[3]{ \det(E_2) }$	
3: $[D, \sim] \leftarrow \text{geigen}(E_1^{-1}, E_2^{-1})$	\triangleright Calcul des valeurs propres généralisées
4: $[\lambda_1, \lambda_2, \lambda_3] \leftarrow \text{trier}(D_{1,1}, D_{2,2}, D_{3,3})$	
5: Si $\ \lambda_1 - \lambda_2\ < \ \lambda_2 - \lambda_3\ $ Alors	\triangleright Recherche de Δ_1 , conique de rang 1
6: $\lambda_D \leftarrow (\lambda_2 + \lambda_1)/2$	
7: Sinon	
8: $\lambda_D \leftarrow (\lambda_2 + \lambda_3)/2$	
9: Fin Si	
10: $\Delta_1 \leftarrow E_1^{-1} - \lambda_D E_2^{-1}$	
11: $[\sim, \sim, V] \leftarrow \text{svd}(\Delta_1)$	
12: $\mathbf{x}_c \leftarrow V_1$	\triangleright Calcul de l'image du centre
13: $\mathbf{v}_\infty \leftarrow E_2 V_1$	\triangleright Calcul de la ligne de fuite

Kim *et al.* [40] proposent également une autre méthode pour pallier la difficulté de calcul et de recherche des valeurs propres. Pour cela, il démontre que la droite passant par le centre des ellipses contient également l'image du centre des cercles. Cette propriété est intéressante et permet de calculer l'emplacement de l'image du centre sur cette droite. Pour cela les auteurs proposent d'utiliser l'invariance du rapport anharmonique en le calculant à partir des points d'intersection entre cette droite et les ellipses. Ces points sont notés $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1, \mathbf{b}_2$ et nous pouvons les observer dans le schéma de la FIGURE 2.6. La position de ces points est connue autant sur le plan de support que sur le plan image et nous pouvons ainsi calculer le rapport anharmonique (voir annexe A.7.1) pour retrouver le point \mathbf{x}_c sur le plan image. Une fois l'image du centre retrouvé \mathbf{x}_c , nous pouvons estimer la rectification affine en procédant de la même manière que précédemment. En effet, il suffit d'utiliser le fait que ce point est en relation polaire avec la ligne de fuite par rapport à l'image des deux cercles, E_1 et E_2 .

Une méthode plus récente proposée par Huang *et al.* fournit une autre interprétation géométrique du calcul de la rectification métrique de l'image de cercles coplanaires [35].

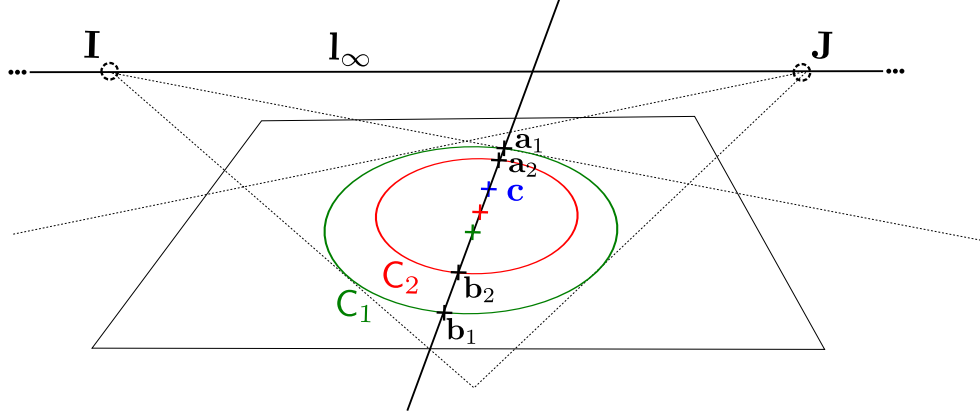


FIGURE 2.6 – Représentation de deux ellipses, images de la projection de deux cercles concentriques. Sur le plan image les points a_1 , a_2 , b_1 et b_2 , sont les points d'intersection des ellipses avec la droite passant par leurs centres. Cette droite contient également le point image du centre des deux cercles.

Les auteurs utilisent la notion de triangle auto-polaire à un cercle. Les auteurs montrent que pour deux cercles concentriques, il y a une infinité de triangles auto-polaires et que tous contiennent un segment sur la ligne de fuite du plan et un sommet qui est l'image du centre. Si nous reprenons la FIGURE 2.4, le triangle en bleu est un exemple particulier d'un triangle auto-polaire complexe formé des deux points cycliques. Les auteurs proposent de retrouver simultanément la ligne de fuite et l'image du centre en recherchant le point et la droite en relation polaire aux deux ellipses. Algébriquement, la solution est cependant équivalente à l'algorithme 1 puisque celle-ci consiste également à calculer les valeurs propres généralisées des matrices E_1 et E_2 .

Calcul préliminaire de la structure euclidienne

Si nous reprenons la conique dégénérée de rang 2 du faisceau, $\tilde{\Delta}_2$ et son image \tilde{D}_2 , celles-ci peuvent également être calculées à partir de l'image de deux cercles concentriques et utilisées pour calculer l'image des points cycliques. Gurdjos *et al.* montrent même qu'il est possible d'étendre ce résultat aux coniques confocales et à leurs images [26]. La conique dégénérée s'obtient de la même façon en utilisant cette fois-ci la valeur propre différente des deux autres. Cette méthode donne cependant des résultats fortement similaires aux autres méthodes.

Calcul d'une homographie après la rectification affine du plan de support

Une fois la rectification affine obtenue c'est-à-dire que la ligne de fuite a été estimée, il est possible d'obtenir une rectification métrique en utilisant une des ellipses. Cependant on peut constater qu'en présence de bruit sur l'image des cercles détectés, les deux ellipses ne sont pas nécessairement l'image de deux cercles parfaitement concentriques. Ainsi selon l'ellipse utilisée pour calculer cette rectification, l'autre ellipse ne sera pas systématiquement rectifiée en cercle. Il est plus commode d'utiliser l'ellipse externe car celle-ci possède plus de points sur l'image et son estimation est ainsi plus précise.

Calvet *et al.* [7] présentent une façon de calculer l'homographie à partir de l'image d'un seul cercle, sa matrice \mathbf{E} , et le vecteur $\mathbf{x}_c = (u, v, 1)^\top$ de l'image du centre.

$$\mathbf{M}_c = \left(\begin{bmatrix} -1 & E_{22}uv & -u \\ 0 & -E_{11}u^2 + 1 & -v \\ -E_{11}u & E_{22}v & 1 \end{bmatrix} \begin{bmatrix} r & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & s \end{bmatrix} \right)^{-1}$$

avec $r = \left(-\frac{E_{22}}{E_{11}}(E_{11}u^2 + E_{22}v^2 + E_{33})\right)^{1/2}$, (2.25)

$$s = \left(-E_{22}(1 - E_{11}u^2)\right)^{1/2}$$

\mathbf{M} représente une homographie de rectification métrique « centrée », c'est-à-dire telle que le centre du cercle corresponde au centre de la rectification. À noter que si l'on impose en plus que le cercle soit de rayon 1 dans le plan de support, cette homographie corrige également l'échelle. L'algorithme 2 présente cette rectification métrique utilisant l'image du centre de l'ellipse $\mathbf{x}_c \sim [x_c \ y_c \ 1]^\top$, et l'ellipse observée fournie sous forme paramétrique (x_0, y_0, a, b, θ) . Les paramètres (x_0, y_0, θ) sont en fait respectivement les coordonnées du centre de l'ellipse dans le plan et l'angle d'inclinaison de son axe principal. Les scalaires a et b représentent respectivement la taille du grand demi-grand et demi-petit axe de l'ellipse.

Algorithm 2 rectifMetrrique_EXc : Homographie à partir de l'image du cercle et de son centre

Entrée: $x_0, y_0, a, b, \theta, x_c, y_c$

Sortie: \mathbf{H}_r

$$\begin{aligned} 1: & \mathbf{T}_{\text{aff}} \leftarrow \begin{bmatrix} \sin \theta & -\cos \theta & -x_0 \sin \theta + y_0 \cos \theta \\ \cos \theta & \sin \theta & -x_0 \cos \theta - y_0 \sin \theta \\ 0 & 0 & 1 \end{bmatrix} \\ 2: & x \leftarrow (x_c - x_0) \sin \theta - (y_c - y_0) \cos \theta \\ 3: & (x_c - x_0) \cos \theta + (y_c - y_0) \sin \theta \\ 4: & k \leftarrow b^2 - x^2 \\ 5: & \phi \leftarrow a^2 b^2 - a^2 x^2 - b^2 y^2 \\ 6: & \mathbf{H}_r \leftarrow \begin{bmatrix} \frac{1}{k\sqrt{\phi}} & 0 & 0 \\ 0 & \frac{1}{k\phi} & 0 \\ 0 & 0 & -\frac{1}{\phi\sqrt{k}} \end{bmatrix} \begin{bmatrix} -1 & 0 & x \\ -xy & -k & b^2 y \\ \frac{xa}{b} & \frac{yb}{a} & -ab \end{bmatrix} \mathbf{T}_{\text{aff}} \end{aligned}$$

2.2.4 À partir d'un cercle, dans le cas calibré

Dans cette section, nous nous intéressons au cas où les paramètres intrinsèques de la caméra sont connus. Dans ce contexte, il est possible de calculer une rectification métrique avec seulement la correspondance d'un cercle et son image. De même que pour le cas précédent, l'homographie ne peut être estimée qu'à une rotation près autour de l'axe du cercle. Replaçons nous ainsi dans les conditions de la section A.9. On rappelle que l'homographie s'écrit :

$$\mathbf{H} \sim \mathbf{K} [\mathbf{R}_1 \ \mathbf{R}_2 \ \mathbf{T}] \text{ avec } \mathbf{R} \in \mathbb{SO}_3^2, \mathbf{T} \in \mathbb{R}^3 \quad (2.26)$$

2. \mathbb{SO}_3 désigne le groupe de rotation 3D.

Si la caméra est calibrée, la matrice \mathbf{K} est connue et l'homographie ne dépend plus que de 6 paramètres conditionnant la pose de la caméra relativement au plan de support, ou 5 paramètres si on se restreint à la pose relative à un cercle du plan.

On peut noter que l'image de la conique absolue, ω , peut être déterminée en utilisant l'équation A.7 de l'annexe A.9. On rappelle que la conique absolue est le lieu de toutes les paires de points cycliques des plans de l'espace projectif 3D et donc également des points cycliques du plan considéré. Un intérêt de connaître la conique absolue est que l'on peut constater que lorsque nous souhaitons obtenir la rectification métrique, nous n'avons en fait besoin que de calculer la ligne de fuite. En effet, les points cycliques sont les points d'intersection de la droite à l'infini et de la conique absolue. L'incidence étant conservée par la projection, l'image des points cycliques est donc l'intersection de la ligne de fuite (image de la droite à l'infini) avec l'image de la conique absolue. Il suffit donc de retrouver la ligne de fuite pour obtenir la rectification métrique dans le cas calibré.

Sans perte de généralité on se place dans le repère 3D tel que le plan a pour équation $z=0$, le vecteur de la normale, \mathbf{N} , au plan de support du marqueur dans le repère cartésien 3D de la caméra respecte donc :

$$\mathbf{N} = \mathbf{Re}_3 = \mathbf{R}_3 \quad (2.27)$$

Ce vecteur peut être obtenu à partir de la ligne de fuite et *vice versa* dans le cas calibré :

$$\mathbf{v}_\infty \sim \mathbf{K}^{-\top} \mathbf{N} \quad (2.28)$$

Ainsi obtenir la normale ou la ligne de fuite, permet de rectifier métriquement l'image d'un plan.

Si on applique la matrice inverse des paramètres intrinsèques, \mathbf{K}^{-1} à tous les points de l'image, on obtient une nouvelle matrice : $\mathbf{Q} = \mathbf{K}^\top \mathbf{E} \mathbf{K}$. On se retrouve en fait dans les mêmes conditions que lorsqu'on utilise l'approximation sur la focale avec $f = 1$ dans la méthode de Chen *et al.* [10] présentée dans la section 2.2.2. En utilisant la décomposition en valeurs propres de $\mathbf{Q} = \mathbf{V}^\top \Sigma \mathbf{V}$, nous avons deux solutions pour la normale, \mathbf{N} qui sont :

$$\mathbf{N} \sim \mathbf{V} \begin{pmatrix} \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} \\ \pm \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \\ 0 \end{pmatrix} \quad (2.29)$$

avec $\mathbf{V} = [\mathbf{V}_1 \mid \mathbf{V}_2 \mid \mathbf{V}_3]$ la matrice concaténant les trois vecteurs propres correspondants, avec chaque \mathbf{V}_i associé à chaque valeur propre λ_i . Ceci nous donne l'algorithme 3 pour le calcul des normales du plan de support à partir de l'image d'un cercle et des paramètres intrinsèques.

On notera qu'une fois la normale obtenue, il est possible de calculer la ligne de fuite avec l'équation 2.28 puis une homographie en utilisant l'algorithme vu plus haut.

2.3 Application à un cas spécifique : le contour d'un bleam

Les méthodes que nous avons présentées au début du chapitre fonctionnent efficacement dans les conditions idéales, lorsque le marqueur possède une assez grande résolution dans l'image. En effet, dans ce cas, l'estimation de l'image des deux cercles peut être réalisée avec un bon degré de précision et l'estimation de l'homographie fonctionne correctement.

2.3. APPLICATION À UN CAS SPÉCIFIQUE : LE CONTOUR D'UN BLEAM

Algorithm 3 `rectifAffine_EK` : Calcul de normale à partir d'un cercle et des paramètres intrinsèques

Entrée: E, K

Sortie: N_0, N_1

- 1: $E \leftarrow \text{signe}(\det(E))E / \sqrt[3]{|\det(E)|}$ ▷ Normalisation de la conique
 - 2: $[[\mu_1, \mu_2, \mu_3], U] \leftarrow \text{eig}(K^{-T}EK^{-1})$
 - 3: $[\lambda_1, \lambda_2, \lambda_3, W] \leftarrow \text{trier}(\mu_1, \mu_2, \mu_3, U)$
 - 4: $N_0 \leftarrow W \begin{bmatrix} \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & +\sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} & 0 \end{bmatrix}^T$ ▷ Calcul des deux normales ambiguës
 - 5: $N_1 \leftarrow W \begin{bmatrix} \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & -\sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} & 0 \end{bmatrix}^T$
-



FIGURE 2.7 – Lecture d'un marqueur bleam loin d'un marqueur

Malheureusement en pratique pour les besoins industriels, les marqueurs sont photographiés loin des caméras, et leurs images sont résolues bien plus faiblement que dans des conditions idéales. La FIGURE 2.7 illustre une lecture de *bleam* dans ces conditions. La distance d'observation ou une forte inclinaison du plan de support du marqueur rendent difficile l'estimation précise de ses paramètres et cela influe sur la précision de l'homographie calculée. De même si dans certains cas, l'hypothèse d'une caméra bien calibrée [63] peut être utilisée, dans les faits, cette situation est très rare. Le problème est qu'en pratique, les caméras sont rarement correctement calibrées car le processus de calibrage est compliqué à mettre en place dans beaucoup de situations concrètes. Sur les appareils les plus récents, des métadonnées peuvent être disponibles et renseignent parfois la distance focale mais elles ne sont pas toujours très précises et parfois incomplètes.

2.3.1 Présentation des méthodes comparées

Dans la suite de cette section, nous allons comparer ces méthodes pour le calcul de la rectification métrique. Nous allons surtout nous concentrer sur deux méthodes principales : celle utilisant les deux cercles concentriques proposée par Kim *et al.* [26] et celle reposant sur l'utilisation des paramètres intrinsèques proposée par Chen *et al.* [11]. Cette étude permettra d'évaluer ces méthodes dans le cas de notre marqueur dans les conditions mentionnées plus tôt, en faisant varier la distance du marqueur à la caméra.

Nous détaillons les méthodes utilisées dans le tableau 2.1 ainsi que le pipeline et les algorithmes suivis dans le schéma de la FIGURE 2.8. Les deux méthodes principales que nous voulons comparer *EE* et *EK* reposent sur les algorithmes *rectifAffine_EK* (voir l'algorithme 3) et *rectifAffine_EE* (voir l'algorithme 1) que nous avons décrits dans la section précédente. Leur principale différence tient dans les données utilisées pour le calcul de l'image du centre, la première utilise les paramètres intrinsèques (K) et la seconde utilise l'image du deuxième cercle (E_2), les deux utilisant principalement l'image du cercle contour (E_1).

CHAPITRE 2. CALCUL DE L'HOMOGRAPHIE POUR LES MARQUEURS CIRCULAIRES

Méthodes	Principe	Données nécessaires
EE	Méthode inspirée par Kim <i>et al.</i> utilisant l'image des deux cercles concentriques en calculant les valeurs propres généralisées de leurs matrices respectives.	Images des deux cercles concentriques
EK	Méthode inspirée par Chen <i>et al.</i> utilisant les paramètres intrinsèques et l'image du cercle externe	Image du cercle externe, paramètres intrinsèques de la caméra
E	Estimation d'une transformation purement affine	Image du cercle externe
VT	Vérité terrain estimée à partir du damier	Image des points du damier

TABLE 2.1 – Méthodes comparées

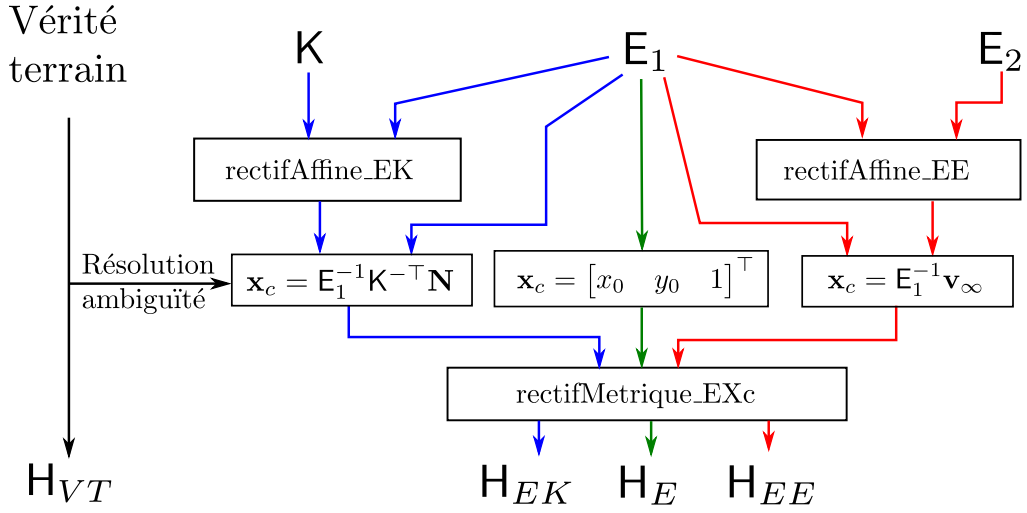


FIGURE 2.8 – Algorithmes de rectification métrique comparés avec le marqueur

Cette comparaison a pour but de montrer que lorsque le marqueur est situé loin de la caméra, la précision de la seconde ellipse utilisée pour la rectification affine peut être insuffisante pour obtenir une rectification correcte. Cependant comme nous l'avons dit précédemment les paramètres intrinsèques ne sont pas non plus toujours bien connus. Ainsi cette comparaison nous permettra de vérifier quelle méthode offre la meilleure piste pour améliorer la rectification métrique.

Rectification métrique à partir de l'image d'un cercle et de son centre

On peut diviser les méthodes en deux étapes, le calcul de la structure affine et la mise à jour euclidienne de cette structure affine de l'homographie de rectification métrique. Pour calculer la rectification métrique, les deux méthodes que nous comparons utilisent l'algorithme 2. Celui-ci reprend la formulation de l'équation 2.25 dans une version légèrement modifiée. Cette modification vient du fait que l'on cherche à obtenir une homographie qui dépend des paramètres de l'ellipse sous la forme paramétrique $\Theta = (x_0, y_0, a, b, \theta)$ plutôt que sous la forme matricielle. L'intérêt est qu'il est plus simple par la suite de paramétrer la solution en fonction des paramètres de l'ellipse

comme nous le verrons par la suite. (x_0, y_0) sont les coordonnées du centre de l'ellipse, a, b sont respectivement le demi-grand et demi-petit axe de l'ellipse et enfin, θ est l'angle de rotation de l'ellipse par rapport à l'axe Ox du plan image et compris entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$. On peut noter que si le repère de notre plan est placé au centre du cercle et que celui-ci est de rayon 1, l'homographie obtenue rectifie l'image du marqueur à une simple rotation près.

Méthode EK

Pour modéliser une caméra incorrectement calibrée, nous définissons une matrice de paramètres intrinsèques par défaut, $K = K_{default}$.

$$K_{default} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.30)$$

avec f qui vaut neuf dixième de la diagonale de l'image, une valeur généralement constatée sur certains *smartphones*, et c_x et c_y , respectivement la demi-longueur et la demi-hauteur de l'image. On fait donc l'hypothèse d'un point principal centré sur l'image et une estimation de la focale approximative. La valeur choisie pour f sera discutée plus amplement dans le chapitre 3, nous conservons cette valeur approchée pour le moment. L'ellipse correspondant à l'image du cercle externe a pour matrice E_1 . Celle-ci est détectée en utilisant un algorithme s'inspirant de la méthode de Fornaciari *et al.* présentée dans la section 1.3.2. Pour l'estimation de ces paramètres, nous utilisons la méthode de Szpak *et al.* présentée dans l'annexe B.3.

On sait que cette méthode donne deux solutions pour la normale du plan de support. On propose dans ce chapitre de supposer cette ambiguïté résolue. Les conditions de résolution de cette ambiguïté seront l'objet du chapitre 4. Dans ce chapitre, nous utilisons la vérité terrain pour la résoudre, on sélectionnera donc la normale la plus proche, en termes d'angle, de la normale vérité terrain.

Méthode EE

La seconde méthode que nous notons *EE* se base sur l'algorithme 1. Cette méthode utilise l'image des deux cercles, E_1 et E_2 . Ces ellipses sont détectées de la même façon que dans la méthode décrite précédemment.

Normalement les valeurs propres généralisées des matrices des coniques E_1^{-1} et E_2^{-1} sont supposées être de la forme $(\lambda_1, \lambda_2, \lambda_2)$. En pratique, à cause de la présence de bruit sur l'image ou de problème de rasterisation, il n'y a pas de valeurs propres de multiplicité 2, mais deux valeurs propres très proches qu'on notera λ_{21} et λ_{22} , il est possible de prendre la moyenne des deux valeurs propres pour obtenir une solution comme proposée dans l'algorithme 1. Ainsi cette méthode mitige l'erreur et ne peut donc garantir que la rectification corrige bien les deux ellipses en cercles. En utilisant, pour calculer la rectification métrique, l'image du premier cercle, on s'assure uniquement que le contour externe est rectifié en cercle. On peut noter également que le rayon relatif entre les deux cercles a une importance sur la rectification estimée. Kim *et al.* [40] montrent que si le cercle interne a un rayon relatif trop grand par rapport au cercle externe, le problème se dégrade et l'erreur de reprojection augmente. Au contraire si le rayon est trop faible la détection de son image dans la photographie devient plus difficile et pose des problèmes de précision. Les auteurs constatent que le rapport idéal entre les rayons est de 0.5,

on observe toutefois qu'il n'y a pas de changement considérable sur l'erreur en dessous d'un rapport de 0.8. Un rapport de 0.75 a été choisi parce qu'il permet d'intégrer plus facilement un identifiant à l'intérieur du cercle interne.

Méthodes de référence

Les autres méthodes sont utilisées comme références. La méthode que nous notons E , consiste à utiliser une seule ellipse et à ne calculer que la transformation affine. Il ne s'agit donc pas à proprement parler d'une rectification métrique. En fait lorsque nous remplaçons les coordonnées de l'image du centre \mathbf{x}_c par les coordonnées du centre de l'ellipse, (x_0, y_0) dans l'algorithme 2, cela revient à calculer une transformation affine qui transforme l'ellipse en un cercle centré et de rayon 1. Ceci permet de comparer l'erreur qu'on obtient pour la rectification métrique avec une transformation qui ne modifie pas la perspective afin de mesurer l'apport de la rectification perspective calculée par les algorithmes.

La dernière rectification, VT est la vérité terrain. Pour la calculer, nous utilisons un damier au centre duquel on dispose le marqueur comme représenté sur la FIGURE 2.9. L'hypothèse est que la mire que constitue ce damier est suffisamment grande pour obtenir une précision de mesure suffisante pour notre travail. Nous faisons ainsi l'hypothèse que l'erreur obtenue par nos méthodes sera bien supérieure aux erreurs obtenues en utilisant le damier.

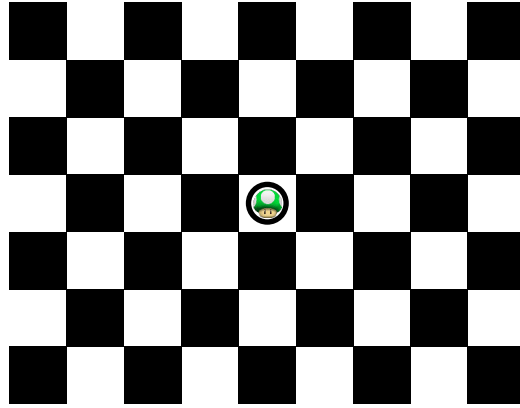


FIGURE 2.9 – Exemple de marqueur avec un damier pour la vérité terrain

Pour obtenir la vérité terrain, nous avons utilisé *OpenCV*, l'algorithme consiste à utiliser les correspondances des points du damier sur plusieurs images pour calibrer la caméra en utilisant la méthode proposée par Zhang [83]. Nous avons donc des paramètres intrinsèques supposés proches d'une modélisation parfaite que nous noterons, $K_{\text{calibré}}$. Pour chaque image, on peut calculer la pose de chaque caméra, ce qui donne la normale du plan de support et également l'homographie qui rectifie le damier dans son plan de support, H_{VT} .

Critère de comparaison

Nous proposons d'utiliser l'erreur de reprojection des points du damier qui nous sert de vérité terrain pour l'homographie (cf. FIGURE 2.9). La grille présentée est composée

Algorithm 4 Calcul de la similitude 2D entre des points

Entrée: $\mathbf{X}, \mathbf{X}', H_r$
Sortie: S

```

1:  $\mathbf{y} \leftarrow H_r \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix}$ 
2:  $\mathbf{Y}_i \leftarrow \mathbf{y}_i / Y_{i,3}$ 
3:  $\bar{\mathbf{X}}'_i = \mathbf{X}_i - \text{mean}(\mathbf{X}')$ 
4:  $\bar{\mathbf{Y}}_i = \mathbf{Y}_i - \text{mean}(\mathbf{Y})$ 
5:  $s_x \leftarrow \sum_{i=1}^{n \times m} \frac{\|\mathbf{X}_i - \bar{\mathbf{X}}'_i\|}{n \times m}$ 
6:  $s_y \leftarrow \sum_{i=1}^{n \times m} \frac{\|\mathbf{Y}_i - \bar{\mathbf{Y}}_i\|}{n \times m}$ 
7:  $\bar{\mathbf{X}}'_i \leftarrow \frac{\bar{\mathbf{X}}'_i}{s_x}$ 
8:  $\bar{\mathbf{Y}}_i \leftarrow \frac{\bar{\mathbf{Y}}_i}{s_y}$ 
9:  $\mathbf{A} \leftarrow \bar{\mathbf{X}}'^T \bar{\mathbf{Y}}$ 
10:  $[\mathbf{U}, \mathbf{S}, \mathbf{V}] \leftarrow \text{svd}(\mathbf{A})$ 
11:  $\mathbf{R} \leftarrow \mathbf{U}\mathbf{V}^T$ 
12:  $u_0 \leftarrow \text{atan2}(\mathbf{R}_{1,2}, \mathbf{R}_{1,1})$ 
13:  $\mathbf{T}_0 \leftarrow \frac{s_x}{s_y} \begin{bmatrix} \cos u_0 & -\sin u_0 \\ \sin u_0 & \cos u_0 \end{bmatrix} \text{mean}(\mathbf{X}') + \text{mean}(\mathbf{Y})$ 
14:  $\Theta \leftarrow \min_{\{u, \mathbf{T}, s\}} \sum_{i=1}^{n \times m} \|\mathbf{X}'_i - s \begin{bmatrix} \cos u & -\sin u \\ \sin u & \cos u \end{bmatrix} \mathbf{Y}_i - \mathbf{T}\| \quad \triangleright \text{Initialisé à } \{u_0, \mathbf{T}_0, \frac{s_x}{s_y}\}$ 
    
```

de $n = 8$ fois $m = 6$ points détectables à l'intersection de chaque case. Ces points détectables dans l'image et de position connue sur le plan de support permettent de mesurer l'erreur sur l'homographie de rectification estimée.

Cependant dans notre cas nous évaluons seulement la rectification métrique. Ainsi nous voulons mesurer la norme des carrés des résidus entre l'image des points du damier et leur reprojection en utilisant la rectification métrique et une similitude inconnue :

$$\sum_{i=1}^{n \times m} \|\mathbf{X}_i - \mathcal{H}(\mathbf{X}'_i)\|^2 = \min_{\mathcal{S}} \sum_{i=1}^{n \times m} \|\mathbf{X}_i - \mathcal{S}(\mathcal{H}_r(\mathbf{X}'_i))\|^2 \quad (2.31)$$

où \mathcal{S} est une similitude plane et \mathcal{H}_r une homographie de rectification métrique.

Pour évaluer cette mesure il est donc nécessaire de calculer la similitude plane entre l'image rectifiée métriquement et le modèle d'apparence du marqueur. Cette tâche peut être réalisée en utilisant les correspondances entre les images des points du damier et les points attendus. Nous fournissons dans l'algorithme 4 l'approche qui peut être utilisée pour calculer la similitude à partir des correspondances entre les points du damier.

L'algorithme 4 donne une approche pour calculer la similitude à partir des correspondances entre les points du damier. Une première solution algébrique est calculée en utilisant une résolution linéaire. Nous posons ensuite le problème de façon géométrique (distance euclidienne) en optimisant l'équation 2.31. Pour cela on utilise l'algorithme des régions de confiance [53] implémenté dans *Matlab*.

2.3.2 Résultats des tests sur la rectification

Pour comparer les méthodes, nous avons pris des photos du damier de la FIGURE 2.9 avec un *smartphone* à des distances différentes du damier. Puis nous avons calculé l'erreur de reprojection pour chaque rectification calculée avec les différentes méthodes :

EE , EK avec les paramètres calibrés avec la vérité terrain, EK avec des paramètres par défaut, E et VT . La FIGURE 2.10 montre les résultats de cette comparaison. On peut constater que la méthode EE fournit en moyenne une erreur de reprojection importante. Cela signifie que la détection de l'ellipse interne n'est pas suffisamment précise pour être utilisée dans le calcul de la ligne de fuite. Le résultat surprenant est que, même sans rectification perspective (méthode E), lorsque l'ellipse est observée loin par rapport au repère caméra, nous obtenons une plus petite erreur de reprojection pour les points du damier. Ceci signifie bien que la ligne de fuite et la perspective ne sont pas retrouvées avec la méthode EE dans ces conditions. Au contraire la méthode EK semble fournir de bien meilleurs résultats et lorsqu'on utilise les paramètres intrinsèques calibrés, la méthode fournit une erreur comparable à la rectification calculée à partir du damier. Sans surprise lorsqu'on utilise cette fois les paramètres par défaut, l'erreur de reprojection augmente, celle-ci n'augmente cependant pas de façon significative et reste bien inférieure à l'erreur moyenne de la méthode E .

Si nous distinguons maintenant les résultats en fonction de la taille apparente de l'ellipse externe correspondant au contour du marqueur, on obtient la FIGURE 2.11. Cette figure permet de mieux comprendre l'observation sur l'erreur pour la méthode EE . En effet lorsque la longueur du demi grand axe de l'ellipse contour diminue dans l'image, la précision sur son estimation est d'autant plus faible car l'influence du bruit et de la discrétisation sur la mesure des points augmente. Ainsi on constate une très grande erreur de précision et l'erreur de reprojection devient très importante, car la perspective n'est plus correctement rectifiée, la FIGURE 2.12 illustre ce problème. On constate en effet que la méthode affine ne déforme pas la perspective alors que la méthode EE amène à une déformation importante. Inversement lorsque celle-ci est plus grande dans l'image, la détection gagne en précision et la méthode EE parvient à de meilleurs résultats que la méthode affine E .

Un autre constat est que la méthode EK avec des paramètres par défaut se comporte très bien quand le marqueur est observé loin de la caméra et celle-ci fournit des résultats proches de la méthode où les paramètres sont calibrés.

2.4 Conclusion sur les méthodes et perspectives

Dans ce chapitre, nous avons présenté les méthodes de calcul de l'homographie à partir de différentes primitives et en particulier, de cercles. Nous avons montré que dans le cas d'une perspective où l'on dispose de l'image d'un cercle, il est suffisant pour calculer l'homographie à une rotation près, c'est-à-dire la rectification métrique centrée et à l'échelle, d'obtenir l'image de la ligne de fuite. Nous avons présenté en détails deux méthodes de l'état de l'art et nous avons proposé de les adapter à notre problème. L'une que nous avons nommée EE utilise l'image d'un second cercle concentrique, l'autre, nommée EK , utilise l'image de la conique absolue calculée à partir des paramètres intrinsèques de la caméra. Les résultats que nous avons obtenus ont permis de mettre en évidence qu'utiliser le cercle interne directement pour le calcul de la perspective dans le cadre de notre marqueur n'apporte pas de résultats significativement bons. Il semble en effet préférable d'utiliser la méthode de Chen *et al.*[10] que nous avons modifiée pour prendre en compte des paramètres intrinsèques quelconques et cela même avec des paramètres intrinsèques approximatifs. Ce constat nous interroge alors sur le rôle des paramètres intrinsèques dans la méthode de rectification que nous avons utilisée et nous pensons qu'il est possible d'améliorer encore ce résultat, notamment en utilisant une information photométrique. La prise en compte de cette information pourrait également

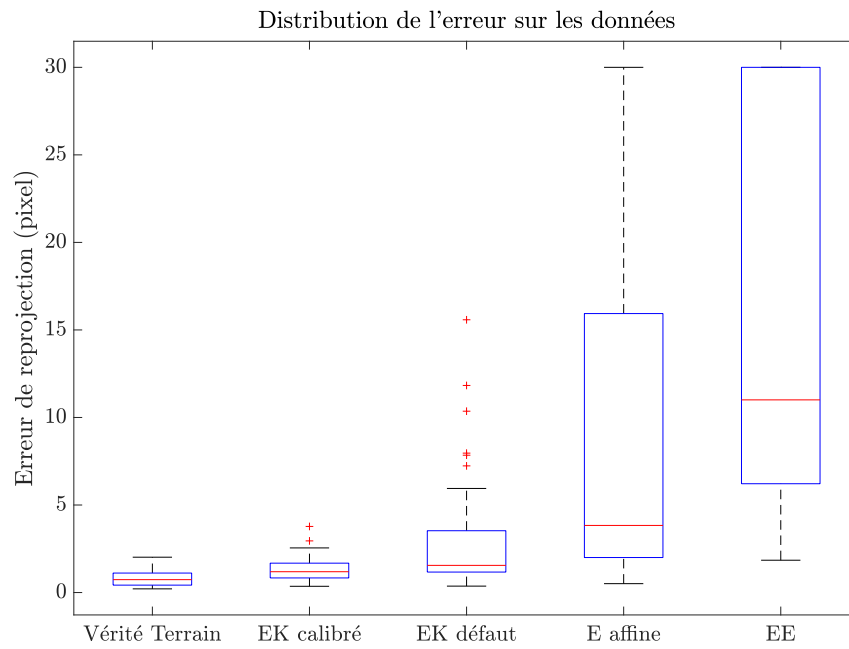


FIGURE 2.10 – Comparaison de la méthode calibrée, EK , et de la méthode à deux cercles concentriques, EE

être utile dans le cas où le capteur est assez loin du marqueur et où la détection de l'ellipse interne dans l'image devient une opération coûteuse et peu précise. C'est ce que nous allons étudier dans le chapitre suivant.

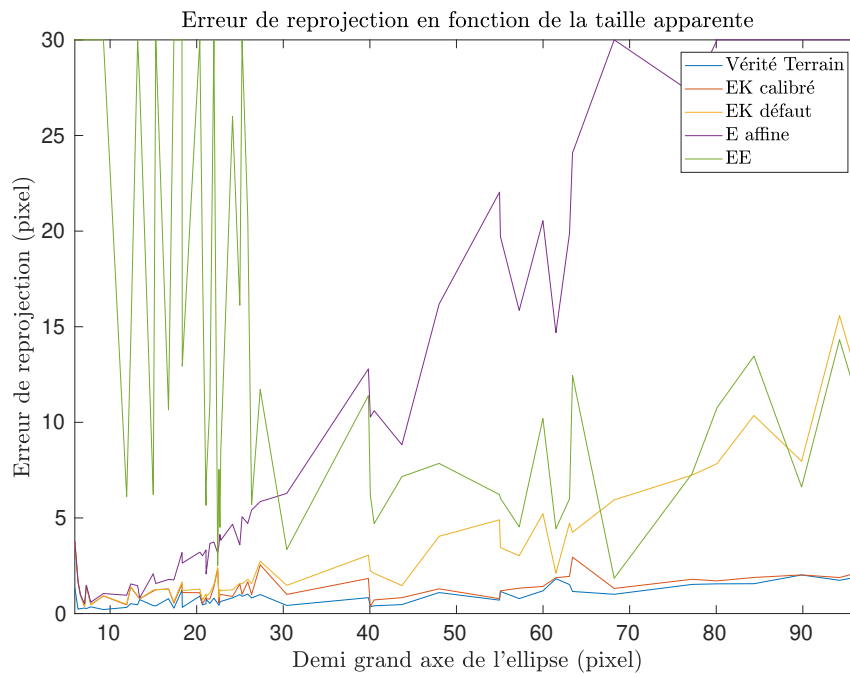


FIGURE 2.11 – Comparaison des méthodes selon la taille de l'ellipse dans l'image

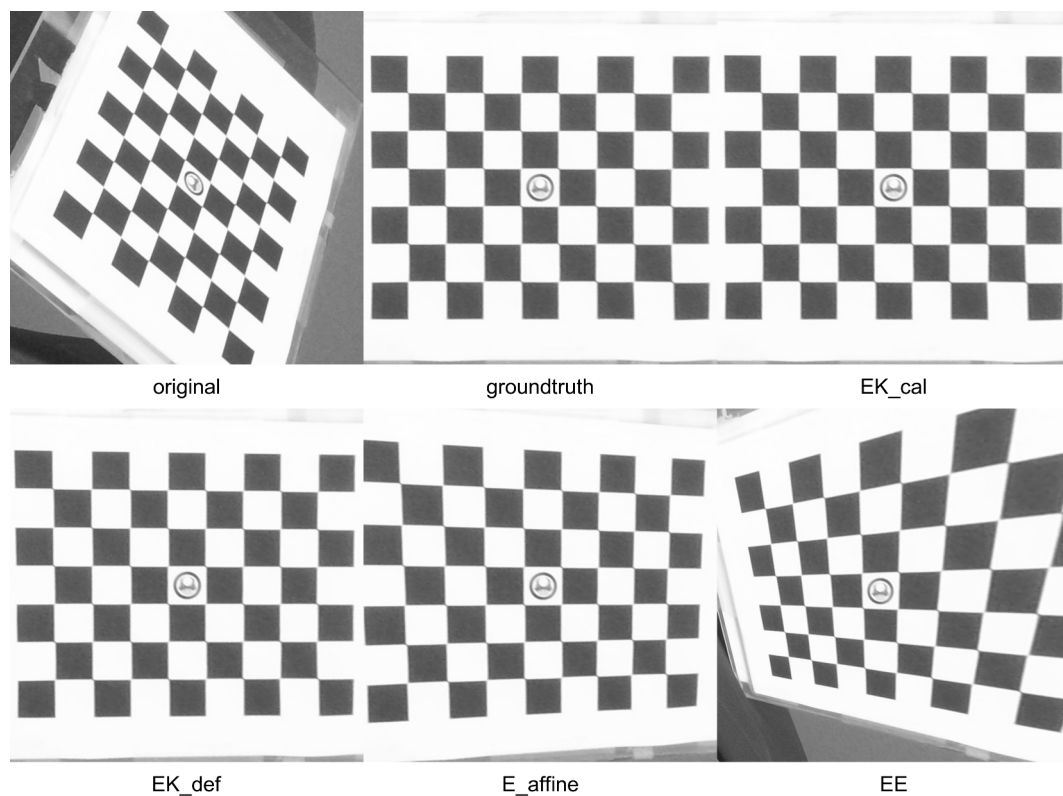


FIGURE 2.12 – Exemple de rectification avec les différentes méthodes

Chapitre 3

Rectification métrique d'un bleam vu par une caméra calibrée par défaut

Sommaire

3.1	Estimation statistique de l'image du centre	46
3.2	Méthode différentiable proposée	47
3.2.1	Fonctions de calcul de l'image du centre	47
3.2.2	Calcul de la normale du plan de support	48
3.2.3	Calcul des valeurs propres de la matrice de l'image du cercle .	49
3.3	Propagation de l'erreur et influence de la focale	50
3.3.1	Propagation d'incertitude sur l'image du centre	50
3.3.2	Incetitude sur les paramètres d'entrée	51
3.3.3	Calcul de la fonction de répartition	51
3.4	Un critère de validation photométrique des candidats . . .	52
3.4.1	Calcul de la similitude d'image	53
3.4.2	Similitude d'image par la mesure statistique des intensités . . .	54
3.4.3	Recherche d'un critère invariant en rotation	56
3.4.4	Création de critères photométriques pour notre méthode	57
3.5	Une évaluation de notre méthode	59
3.5.1	Construction d'un calibrage par défaut	59
3.5.2	Évaluation des candidats générés	61
3.6	Conclusion et perspectives sur la méthode proposée	62

3.1 Estimation statistique de l'image du centre

Pour calculer une rectification métrique d'un marqueur circulaire par une approche géométrique, nous avons mis en évidence deux méthodes, la méthode *EK*, qui utilise à la fois l'image d'un seul cercle et les paramètres intrinsèques et la méthode *EE* qui utilise les images des deux cercles concentriques. Nous avons constaté que les deux méthodes ne se différencient que par le calcul des propriétés affines du plan. Dans le cas de l'image d'un cercle, la ligne de fuite est en relation directe (liée à la polarité) avec l'image de son centre. Autrement dit, étant donné le contour, les deux données sont équivalentes. Ainsi dans ce cas, le calcul des propriétés affines revient à obtenir l'image du centre du cercle. Ce que nous observons dans les résultats du chapitre 2 c'est que la précision d'estimation sur les coordonnées du vecteur de l'image du centre affecte de façon significative l'erreur de reprojection obtenue. Ainsi, lorsque les paramètres intrinsèques sont correctement calibrés, nous approchons la précision obtenue avec un damier pourtant 8 fois plus grand dans l'image. D'autre part, nous constatons que lorsque la taille relative du marqueur dans l'image est faible, la méthode *EK*, même avec un calibrage approximatif, donne une précision très convenable. Ces constats nous amènent à nous poser la question suivante : Est-il possible de trouver l'image du centre qui minimise l'erreur de reprojection, même lorsque nous ne possédons que des *a priori*¹ sur le calibrage ?

Afin de répondre à cette question, nous proposons d'introduire une approche fondée sur la génération de candidats. L'idée est d'utiliser ce que nous appelons des *a priori* sur les paramètres intrinsèques et l'observation de l'image du cercle pour calculer des candidats potentiels pour l'image du centre du contour circulaire. Cet ensemble de candidats décrit ainsi un ensemble de vecteurs de coordonnées probables pour l'image du centre. Dans l'hypothèse où cet ensemble contiendrait effectivement un vecteur proche de la vérité terrain, c'est-à-dire les coordonnées réelles de l'image du centre, que nous notons \mathbf{x}_c , nous pouvons supposer qu'il est possible de l'extraire et donc d'obtenir une rectification plus précise. En utilisant le modèle d'apparence du marqueur étudié, il est possible de définir un critère photométrique, c'est-à-dire un critère qui, à partir de la mesure des intensités de l'image, permet d'identifier la transformation rectifiant correctement la vue du marqueur prise par la caméra au modèle imprimé sur le plan de support. Un tel critère doit donc permettre de sélectionner dans notre ensemble de candidats, le « candidat idéal », c'est-à-dire le vecteur de coordonnées offrant une rectification métrique qui minimise l'erreur de reprojection du marqueur.

Nous proposons ainsi de suivre le schéma suivant :

1. Définir un calibrage par défaut² représentant les *a priori* sur les paramètres de calibrage ;
2. Détecter l'image du contour circulaire du marqueur ;
3. Générer des candidats pour l'image du centre en utilisant le calibrage par défaut et l'image du cercle ;
4. Sélectionner parmi les candidats ceux donnant la meilleure rectification au sens d'un certain critère photométrique.

Pour réaliser l'étape 3, il serait intéressant d'obtenir une formulation différentiable qui donne l'image du centre en fonction des paramètres intrinsèques et de l'image du

1. Nous nommons ici *a priori* le fait d'avoir un calibrage par défaut ou très approximatif.

2. Ce terme désigne le fait d'attribuer des valeurs par défaut aux paramètres du calibrage, sans utiliser une méthode de calibrage. Nous reviendrons sur cette notion dans la suite de ce chapitre.

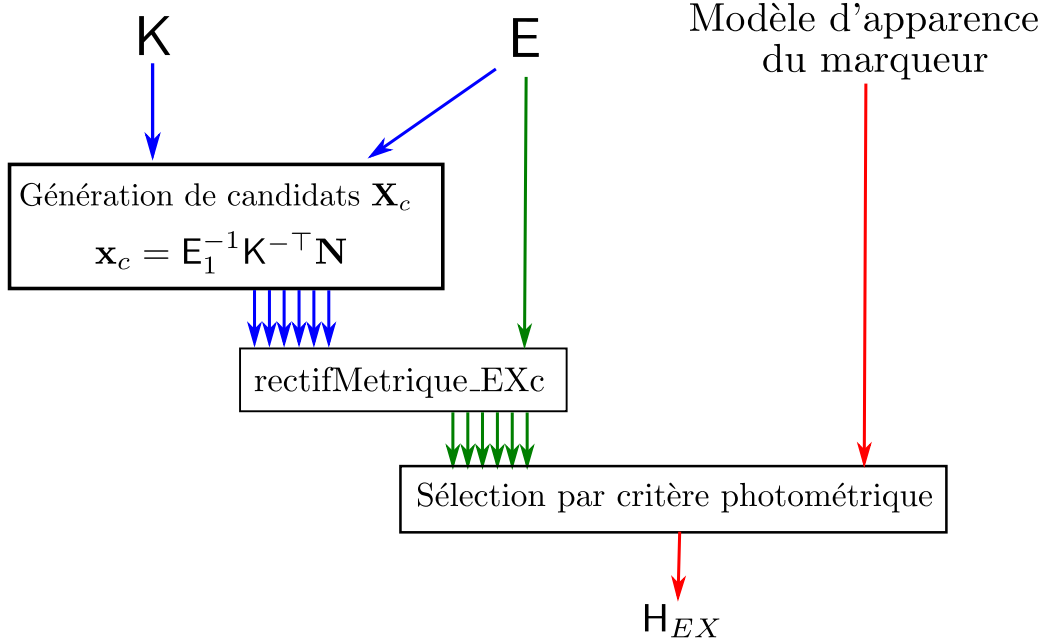


FIGURE 3.1 – Schéma des étapes de la méthode de rectification proposée.

cercle. En effet, l'intérêt d'une telle formulation est qu'il est possible de propager l'incertitude sur les paramètres d'entrée provenant à la fois de l'estimation de l'ellipse externe et des connaissances approximatives des paramètres intrinsèques, ce que nous avons appelé précédemment le modèle de caméra par défaut. La FIGURE 3.1 présente les étapes 3 et 4 de la méthode proposée dans ce chapitre. Dans un premier temps, nous recherchons donc une expression différentiable de l'image du centre. Puis, nous examinons comment il est possible de construire un modèle de calibrage par défaut et de l'utiliser pour générer des candidats pour la rectification métrique. Nous discuterons ensuite des critères photométriques pour sélectionner le candidat idéal. Enfin, nous présenterons les résultats obtenus.

3.2 Méthode différentiable proposée

3.2.1 Fonctions de calcul de l'image du centre

Nous voulons donner les coordonnées de l'image du centre sous une forme analytique qui est différentiable, ce qui requiert de n'utiliser ni de « si-alors-sinon » ni de $\max()$ ou $\min()$, contrairement aux définitions précédentes [5, 11], comme dans l'algorithme 3, où un tri est requis. Cette nouvelle formulation nous permet ainsi d'obtenir une forme dérivable tout en restant équivalente aux formulations présentées dans le chapitre 2.

Nous appelons \mathcal{F} , la fonction de calcul de l'image du centre. Cette fonction a pour variables les éléments de la matrice \mathbf{K} , matrice des paramètres intrinsèques de la caméra, et les paramètres de l'ellipse observée, $\Theta = (x_0, y_0, a, b, \theta)$ (cf. section 2.3.1), qu'on donne ici sous forme paramétrique et qui correspond à l'image du cercle. Nous définissons la matrice des paramètres intrinsèques comme dans l'équation (A.1) de l'annexe A.2. C'est-à-dire que celle-ci est paramétrée par cinq scalaires : $(f_x, f_y, \alpha, u_0, v_0)$. Le but est donc d'obtenir une expression différentiable du vecteur de coordonnées cartésiennes, \mathbf{X}_c , correspondant à l'image du centre du marqueur à partir des paramètres

scalaires $(x_0, y_0, a, b, \theta, f_x, f_y, \alpha, u_0, v_0)$. On définit ainsi la fonction \mathcal{F} comme une fonction vectorielle qui retourne le vecteur de l'image du centre en coordonnées cartésiennes $\mathbf{X}_c = (x_c, y_c)$ en fonction des paramètres scalaires. Puisqu'il existe deux solutions ambigües, on définit deux fonctions \mathcal{F}_0 et \mathcal{F}_1 qui donnent les deux images du centre.

$$\begin{aligned}\mathbf{X}_{c,0} &= \mathcal{F}_0(x_0, y_0, a, b, \theta, f_x, f_y, \alpha, u_0, v_0) \\ \mathbf{X}_{c,1} &= \mathcal{F}_1(x_0, y_0, a, b, \theta, f_x, f_y, \alpha, u_0, v_0)\end{aligned}$$

avec le vecteur de coordonnées homogènes, \mathbf{x}_c défini comme :

$$\begin{bmatrix} \mathbf{X}_c \\ 1 \end{bmatrix} \sim \mathbf{x}_c \quad (3.1)$$

Dans cette section, nous allons détailler comment nous pouvons obtenir de telles fonctions. Dans un premier temps, on se place dans le repère normalisé de la caméra où la projection perspective a pour matrice $\mathbf{P} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}$ car $\mathbf{K}' = \mathbf{I}$. Pour cela, on peut appliquer aux points de l'image, la transformation suivante :

$$\mathbf{x}' \sim \mathbf{K}^{-1} \mathbf{x} \quad (3.2)$$

On fait de même pour l'image du cercle, paramétrée par la matrice symétrique, \mathbf{E} :

$$\mathbf{Q} = \mathbf{K}^\top \mathbf{E} \mathbf{K} \quad (3.3)$$

L'ellipse \mathbf{Q} peut être entièrement paramétrée par $(x_0, y_0, a, b, \theta, f_x, f_y, \alpha, u_0, v_0)$:

$$\mathbf{Q} = \mathbf{K}^\top \begin{bmatrix} \cos \theta & \sin \theta & p_x \\ -\sin \theta & \cos \theta & p_y \\ 0 & 0 & 1 \end{bmatrix}^\top \begin{bmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{b^2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta & p_x \\ -\sin \theta & \cos \theta & p_y \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}$$

avec

$$p_x = -x_0 \cos \theta - y_0 \sin \theta \text{ et } p_y = x_0 \sin \theta - y_0 \cos \theta$$

et

$$\mathbf{K} = \begin{bmatrix} f_x & \alpha & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.2.2 Calcul de la normale du plan de support

Nous détaillons la fonction de calcul de la normale du plan de support du cercle dans le repère 3D de la caméra. On peut noter que dans le repère caméra lorsque $\mathbf{K}' = \mathbf{I}$, la normale a les mêmes coordonnées homogènes que la ligne de fuite dans le repère de l'image. Ainsi puisque la ligne de fuite est la droite polaire de l'image du centre du cercle par rapport à l'image du cercle, de matrice \mathbf{Q} , cf. annexe A.7.3 :

$$\mathbf{x}_c \sim \mathbf{K}^{-1} \mathbf{Q}^{-1} \mathbf{N} \quad (3.4)$$

Le problème que nous résolvons ci-dessous est donc équivalent à donner une forme analytique aux deux solutions pour le vecteur, \mathbf{N} , de la normale du plan de support du cercle.

Supposons les faits suivants vérifiés :

— $\mathbf{Q} \in \mathcal{S}_3(\mathbb{R})$ est la matrice de l'image d'un cercle ;

- $\{\lambda_i\}$, $i = 1..3$, est l'ensemble des valeurs propres de \mathbf{Q} et $\{\mathbf{V}_i\}$ celui des vecteurs propres *unitaires* associés, c'est-à-dire que $\|\mathbf{V}_i\| = 1$;
- $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

Théorème 3.2.1. *Les deux solutions pour \mathbf{N} , vecteur unitaire de la normale au plan de support du cercle, sont*

$$\mathbf{N}_{0,1} = \frac{\sqrt{\lambda_1 - \lambda_2} \mathbf{V}_1 \pm \sqrt{\lambda_2 - \lambda_3} \mathbf{V}_3}{\|\sqrt{\lambda_1 - \lambda_2} \mathbf{V}_1 \pm \sqrt{\lambda_2 - \lambda_3} \mathbf{V}_3\|} \quad (3.5)$$

Ce résultat provient directement du chapitre 2.

3.2.3 Calcul des valeurs propres de la matrice de l'image du cercle

Généralement les méthodes pour calculer la normale du plan de support à partir de l'image d'un cercle de ce plan reviennent à calculer les vecteurs propres de la matrice du cercle dans le repère de la caméra normalisée. Ces méthodes reposent ainsi sur la décomposition en valeurs singulières qui a le défaut de ne pas être directement différentiable, bien que des méthodes ont été développées pour estimer la matrice jacobienne de cet opérateur comme celle proposée par Papadopoulos *et al.* [64]. Dans le cas de la matrice symétrique réelle, \mathbf{Q} , il est possible d'obtenir le calcul d'une forme symbolique des valeurs propres ordinaires en s'inspirant des travaux de Kopp *et al.* [43].

Théorème 3.2.2. *Les valeurs propres de \mathbf{Q} sont*

$$\lambda_1 = 2p \cos \frac{\alpha}{3} + q \quad (3.6)$$

$$\lambda_2 = p \left(\sqrt{3} \sin \frac{\alpha}{3} - \cos \frac{\alpha}{3} \right) + q \quad (3.7)$$

$$\lambda_3 = -p \left(\sqrt{3} \sin \frac{\alpha}{3} + \cos \frac{\alpha}{3} \right) + q \quad (3.8)$$

où

$$\alpha = \arccos \left(\frac{1}{2} \det \frac{1}{p} \mathbf{F} \right) \in [0, \pi]$$

$$q = \frac{1}{3} \text{trace}(\mathbf{Q})$$

$$\mathbf{F} = \mathbf{Q} - q\mathbf{I}$$

$$p = \sqrt{\text{trace} \left(\left(\frac{1}{6} (\mathbf{F})^2 \right) \right)}$$

et l'on a bien,

$$\lambda_1 \geq \lambda_2 \geq \lambda_3$$

Dans ce chapitre, nous admettons ce résultat, la démonstration sera faite dans le chapitre 4.

Théorème 3.2.3. *Les formes analytiques de $\lambda_1 - \lambda_2$ et $\lambda_2 - \lambda_3$ dans l'équation (3.5) sont*

$$\lambda_1 - \lambda_2 = \cos \frac{\alpha}{3} - \frac{\sqrt{3}}{3} \sin \frac{\alpha}{3} \quad (3.9)$$

$$\lambda_2 - \lambda_3 = 2 \frac{\sqrt{3}}{3} \sin \frac{\alpha}{3} \quad (3.10)$$

L'intérêt de ce dernier résultat est de montrer, d'une part qu'il est possible de calculer directement $\lambda_1 - \lambda_2$ et $\lambda_2 - \lambda_3$, sans passer par un algorithme de décomposition en valeurs propres, et, d'autre part, que la nécessité de trier les valeurs propres ne se pose plus.

Ensuite, il est assez simple de calculer les vecteurs propres de \mathbf{Q} . Si $\mathbf{D} = \mathbf{Q} - \lambda \mathbf{I}$ est de rang 2 alors le vecteur unitaire $\mathbf{V} \in \mathbb{R}^3$ tel que $\mathbf{D}\mathbf{V} = \mathbf{0}$ est de la forme :

$$\mathbf{V} = \frac{\mathbf{W}}{\|\mathbf{W}\|} \text{ où } \mathbf{W} = \begin{bmatrix} -D_{23}^2 + D_{22}D_{33} \\ D_{13}D_{23} - D_{12}D_{33} \\ D_{12}D_{23} - D_{13}D_{22} \end{bmatrix} \quad (3.11)$$

Nous insistons sur le fait qu'une fois ces vecteurs obtenus, les deux solutions pour l'image du centre se déduisent directement à partir des égalités (3.4) et (3.5). Tous les éléments pour construire les deux fonctions \mathcal{F}_1 et \mathcal{F}_2 sont donc maintenant définis analytiquement.

Nous proposons également dans l'annexe C, une évaluation de cette méthode de rectification métrique. Nous y démontrons expérimentalement la validité de cette méthode pour le calcul de la pose d'un cercle. De plus, nous validons le comportement efficace de la méthode face à une variation de la distance focale en présentant l'influence sur la précision du calcul de la pose.

3.3 Propagation de l'erreur et influence de la focale

3.3.1 Propagation d'incertitude sur l'image du centre

Pour calculer la rectification métrique à partir de l'algorithme 2, nous avons besoin de (x_c, y_c) , les coordonnées cartésiennes de l'image du centre. Nous souhaitons rechercher ce vecteur dans un sous-espace de \mathbb{R}^2 en prenant en compte nos « a priori » sur les paramètres intrinsèques et les paramètres de l'ellipse détectée. Une de nos contributions est de restreindre ce sous-espace à partir de considérations statistiques en introduisant la notion de « calibrage par défaut ». Les contraintes de fabrication sur les caméras intégrées aux smartphones impliquent en effet des contraintes sur la taille des lentilles et des capteurs de ces appareils. Une des répercussions est que les paramètres intrinsèques des smartphones affichent souvent des valeurs similaires ou très semblables.

Notre idée est alors d'interpréter la variance entre les appareils comme une donnée sur l'incertitude d'une caméra possédant des paramètres « par défaut ». Ces considérations statistiques sur les paramètres intrinsèques couplées aux incertitudes des paramètres de l'ellipse détectée comme ellipse externe permettent la restriction de l'espace de recherche. Ayant montré que calculer la rectification métrique de l'image est équivalent à rechercher l'image du centre, nous souhaitons propager ces incertitudes sur l'image du centre estimée avec le calibrage par défaut. Pour cela, nous pouvons utiliser la fonction de calcul de l'image du centre (ou de la normale) différentiable qui a été obtenue dans la section précédente, on rappelle que ce n'est pas le cas pour celles proposées dans la littérature.

On montrera aussi que l'on peut définir une incertitude autour de cette valeur avec une variance très faible autour d'un vecteur moyen, ce qui nous donne une précision subpixelique sur les coordonnées de ce point. On observera les distributions des centres lors de la variation à la fois des paramètres de l'ellipse et des paramètres de la caméra (sur la variation de la focale et des coordonnées du point principal).

3.3.2 Incertitude sur les paramètres d'entrée

Étant donné que la caméra n'est pas calibrée, on va définir un « calibrage par défaut » à partir d'un vecteur aléatoire \mathbf{K} . Nous faisons l'hypothèse très simple que le facteur d'obliquité, α , est nul et que les distances focales pixéliques, f_x et f_y sont égales, nous avons ainsi $f = f_x = f_y$. Ces hypothèses consistent donc à supposer un capteur dont la matrice est constituée de pixels carrés. Les paramètres intrinsèques du vecteur \mathbf{K} sont ainsi réduits à trois paramètres $\mathbf{K} = (f, u_0, v_0)$.

La focale pixélique dépend en grande partie de la largeur et de la hauteur de la photographie. On peut ainsi l'exprimer comme une fonction, produit de la diagonale en pixel de la photographie et d'une variable aléatoire. Les paramètres statistiques de cette variable aléatoire peuvent être estimés de façon heuristique, comme nous l'avons fait dans nos expériences cf. section 3.5.1 à partir des données de calibrage de plusieurs caméras. On peut suivre le même raisonnement pour les coordonnées de l'image du point principal qui sont généralement situées proches du centre de l'image. Ainsi les coordonnées u_0 et v_0 peuvent être exprimées respectivement en tant que produit de la demi-largeur et de la demi-hauteur de l'image avec une variable aléatoire centrée autour de 1 dont les paramètres statistiques peuvent être obtenus heuristiquement. On suppose les paramètres intrinsèques indépendants, en effet en considérant le modèle de caméra simplifié présenté, l'indépendance des paramètres revient à justifier que la taille du capteur photométrique en largeur et en longueur et la position horizontale et verticale de celui-ci par rapport à la lentille sont trois paramètres indépendants. On peut alors attribuer à ces paramètres des écarts types de façon empirique afin d'obtenir une loi de distribution sur \mathbf{K} .

Concernant les paramètres de l'ellipse et leurs incertitudes, nous avons utilisé la méthode d'ajustement des points à une ellipse proposée par Spzak *et al.* [77] et présentée dans l'annexe B.3. Les auteurs proposent une méthode d'estimation à la fois des paramètres de l'ellipse, condensée dans le vecteur Θ , et de son intervalle de confiance à partir de l'hypothèse d'un bruit gaussien identique pour tous les points du contour et de l'estimateur de Sampson. Les étapes pour le calcul de $\bar{\Theta}$ et Σ_{Θ} se résument ainsi :

1. Détection des points contours d'une ellipse ;
2. Calcul d'une première estimation algébrique et calcul de la matrice de covariance associée aux points du contour ;
3. Optimisation et calcul de la matrice de covariance associée aux paramètres de l'ellipse.

3.3.3 Calcul de la fonction de répartition

Nous considérons que le vecteur des coordonnées cartésiennes de l'image du centre, $\mathbf{X}_c = (x_c, y_c)^\top$, est un vecteur aléatoire de moyenne $\bar{\mathbf{X}}_c$ et de matrice de covariance $\Sigma_{\mathbf{X}_c}$. Nous utilisons cette hypothèse pour générer des candidats en définissant une loi de distribution pour \mathbf{X}_c , paramétrée par $\bar{\mathbf{X}}_c$ et $\Sigma_{\mathbf{X}_c}$. Pour trouver $\bar{\mathbf{X}}_c$ et $\Sigma_{\mathbf{X}_c}$, nous considérons que les connaissances dont nous disposons sont aussi associées à des vecteurs aléatoires, à savoir que les paramètres de calibrage « par défaut » sont stockés dans un vecteur aléatoire \mathbf{K} et que les paramètres de l'ellipse détectée comme image du cercle, sont stockés dans un vecteur aléatoire Θ . On considère la fonction \mathcal{F}_i avec $i \in \{0, 1\}$ telle que $\mathbf{X}_{c,i} = \mathcal{F}_i(\mathbf{K}, \Theta)$, où \mathcal{F}_i est la fonction définie dans la section 3.2.1. Sachant que $\bar{\mathbf{X}}_{c,i} = \mathcal{F}_i(\bar{\mathbf{K}}, \bar{\Theta})$, nous utilisons une approximation du premier ordre de la fonction

\mathcal{F}_i pour obtenir une valeur approchée de \mathbf{X}_c ,

$$\Sigma_{\mathbf{X}_c} = \mathbf{J}_{\mathcal{F}_i}(\bar{\mathbf{K}}, \bar{\Theta}) \begin{pmatrix} \Sigma_{\mathbf{K}} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\Theta} \end{pmatrix} \mathbf{J}_{\mathcal{F}_i}(\bar{\mathbf{K}}, \bar{\Theta})^\top \quad (3.12)$$

où $\mathbf{J}_{\mathcal{F}_i}$ est la matrice Jacobienne de \mathcal{F}_i , comme cela est décrit par exemple dans [30, p. 139].

Étant donné qu'il existe deux solutions pour \mathbf{X}_c à partir de \mathbf{K} et Θ , elles sont traitées de façon indépendante ce qui nous conduira dans la partie suivante à posséder deux lois pour l'image du centre et à générer ainsi deux ensembles de candidats. Une hypothèse gaussienne sur la distribution de l'image du centre nous permet d'exprimer simplement la loi de X_i avec $\bar{\mathbf{X}}_{c,i}$ et $\Sigma_{\mathbf{X}_{c,i}}$, nous pouvons visualiser un exemple de cette distribution sur la FIGURE 3.2. On peut y voir l'image d'un marqueur sur l'axe OXY , l'axe Z (la hauteur) correspond à la fonction de densité de probabilité pour le point de coordonnées (x, y) .

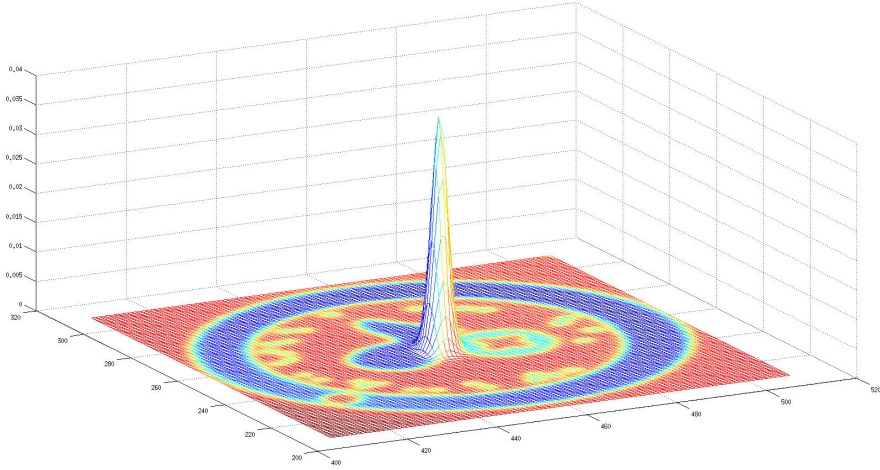


FIGURE 3.2 – Visualisation de la distribution de l'image du centre sur l'image d'un marqueur

Puisqu'a priori les deux cas d'ambiguïté sont équiprobables, nous pouvons construire la loi pour le vecteur \mathbf{X}_c en combinant les lois de probabilité pour $\mathbf{X}_{c,0}$ et $\mathbf{X}_{c,1}$.

3.4 Un critère de validation photométrique des candidats

La connaissance sur la loi du vecteur \mathbf{X}_c nous permet de générer des candidats de façon aléatoire qui respectent les distributions des vecteurs \mathbf{K} et Θ . L'application de l'algorithme 2 avec un candidat et la valeur moyenne, $\bar{\Theta}$, pour l'image du cercle nous permet ensuite d'associer à chaque candidat une homographie de rectification et ainsi une image rectifiée du marqueur notée I^{rect} .

Si on suppose disposer d'un modèle d'apparence du *bleam*, par exemple sous la forme d'une image fronto-parallèle de celui-ci qu'on notera I^{ref} , il est possible de l'utiliser pour rechercher parmi les images rectifiées celles qui s'en approchent le plus.

Pour cela il est nécessaire de définir un critère qui évalue la similitude (ou la dissimilitude) entre deux images du marqueur rectifié. Le but de cette section est d'étudier

un critère pertinent pour mesurer cette différence. Nous étudierons dans un premier temps les critères pertinents de la littérature pour le cas du marqueur *bleam* puis nous retiendrons quelques critères que nous évaluerons.

3.4.1 Calcul de la similitude d'image

Les approches pour mesurer la similitude ou au contraire la dissimilitude entre deux images ont été étudiées abondamment pour résoudre les problèmes de reconnaissance et de recherche d'image ou de calcul de transformation géométrique d'images. Les auteurs montrent que la performance des mesures dépend fortement de la nature des images et de leurs déformations, Goshtasby propose une revue avec comparaison de certaines méthodes [25].

Parmi les mesures classiquement utilisées, on trouve la distance en norme L_2 :

$$d_{L2}(I^{ref}, I^{rect}) = \sum_{x,y} \|I_{x,y}^{ref} - I_{x,y}^{rect}\|^2 \quad (3.13)$$

Ce critère a l'avantage de définir une distance sur l'espace métrique des images, de plus sa convexité facilite son utilisation dans des algorithmes d'optimisation. Puisque cette mesure amplifie des variations d'intensité, elle n'est pas robuste à une image bruitée ou à un changement d'éclairage de la scène, elle est par contre bien adaptée pour mesurer une variation géométrique comme celle due à la perspective que nous voulons corriger. Puisque nous évaluons des candidats et que nous ne mettons pas en place un algorithme d'optimisation, d'autres mesures peuvent être utilisées dans notre cas. La corrélation de Pearson (ou *NCC* pour *Normalized Cross-Correlation*) est également une mesure très populaire :

$$d_{NCC}(I^{ref}, I^{rect}) = \frac{\sum_{x,y} (I_{x,y}^{ref} - \bar{I}^{ref})(I_{x,y}^{rect} - \bar{I}^{rect})}{\sqrt{\sum_{x,y} (I_{x,y}^{ref} - \bar{I}^{ref})^2 \sum_{x,y} (I_{x,y}^{rect} - \bar{I}^{rect})^2}} \quad (3.14)$$

Cette mesure est d'ailleurs utilisée pour un autre marqueur circulaire [63], elle a l'avantage d'être adaptée à un changement affine d'intensité ce qui modélise correctement tout changement d'éclairage lambertien dans une scène 3D. La norme L_2 peut aussi être adaptée pour une transformation affine de l'intensité en centrant les intensités et en les divisant par leurs écarts types [18]. D'autres classes de critères se basent sur la répartition des intensités dans les images, comme celles fondées sur l'énergie de la répartition de probabilité de la loi conjointe [25, p. 22] entre les deux images. Ces mesures ont en effet l'avantage d'être robustes aux transformations non affines de l'intensité, mais elles ont aussi le défaut de ne pas être aussi discriminantes pour les transformations géométriques [25, p. 50].

D'autres mesures comme des mesures fondées sur le rang des valeurs d'intensités dans l'image sont également des mesures utilisées [25, p. 39]. Elles sont toutefois moins adaptées à notre problème dans la mesure où celles-ci demandent un temps d'évaluation trop important. Ces mesures requièrent en effet l'application d'un filtre gaussien pour éviter que deux valeurs d'intensité soient associées à un même rang ainsi que le tri des valeurs d'intensité de toutes les images par ordre croissant pour obtenir leur rang. Dans notre problème, le critère doit pouvoir être évalué rapidement, car il doit être mesuré sur chacun des candidats pour la rectification métrique.

3.4.2 Similitude d'image par la mesure statistique des intensités

Une autre façon de mesurer la similitude entre deux images est d'étudier la distribution statistique des valeurs d'intensité en niveau de gris. Ces méthodes ont l'intérêt de ne pas dépendre de la structure de l'image mais uniquement de la distribution des valeurs d'intensités. Pour cela un histogramme est utilisé pour trier les valeurs en niveau de gris. Les mesures photométriques fondées sur l'histogramme proposent de classer les intensités des pixels de l'image de façon ordonnée dans des classes, appelées *bins*. Chaque élément des vecteurs, P_i , contient le nombre de pixels de chaque classe ainsi quantifiés. Par exemple si n désigne le nombre de *bins*, on peut par exemple construire un histogramme à partir d'intervalles d'intensités suivant une loi uniforme :

$$P_i = \text{card} \left\{ I_{x,y} \text{ tel que } \frac{i-1}{n} \leq \frac{I_{x,y}}{255} \leq \frac{i}{n} \right\} \quad (3.15)$$

On distingue alors les méthodes qui comparent les *bins* un à un des méthodes *cross-bins* croisant les *bins* pour prendre en compte les problèmes d'erreur de quantification. En effet si les intensités sont transformées à cause d'un changement d'éclairage, celles-ci se retrouveront classées dans des *bins* différents.

Les méthodes de comparaisons *bin* à *bin* les plus communes sont les comparaisons en norme L_1 et L_2 qu'on peut généraliser par la distance de Minkowski :

$$d_L(\mathbf{P}, \mathbf{Q}) = \sum_{i=1}^n (|P_i - Q_i|^r)^{\frac{1}{r}}, r \in \mathbb{Z}^* \quad (3.16)$$

À noter que lorsque r tend vers l'infini, on obtient la distance de Kolmogorov-Smirnov qui est très utilisée pour la comparaison de séries unidimensionnelles. Le problème de ces mesures est qu'elles sont sensibles aux *bins* de grandes cardinalités et moins sensibles aux *bins* de petites cardinalités, la mesure du χ^2 permet de contrer cet effet en ajoutant un poids dépendant de la taille du *bin* [68] :

$$d_{\chi^2}(\mathbf{P}, \mathbf{Q}) = \frac{1}{2} \sum_{i=1}^n \frac{(P_i - Q_i)^2}{P_i} \quad (3.17)$$

De même la mesure de Kullback-Leibler, fondée sur la théorie de l'information s'écrit :

$$d_{KL}(\mathbf{P}, \mathbf{Q}) = \sum_{i=1}^n P_i \log\left(\frac{P_i}{Q_i}\right) \quad (3.18)$$

Comme dit précédemment, ces méthodes présentent l'avantage d'être discriminantes au regard de transformations géométriques, mais sont malheureusement sensibles à une mauvaise quantification en cas de variations d'intensité entre \mathbf{P} et \mathbf{Q} dues à un changement d'éclairage.

L'autre classe de mesures de distance entre deux histogrammes est celle des mesures *cross-bin*, celles ci ne prennent pas uniquement en compte la différence de tailles des *bins* de même indice, ce qui permet d'être plus tolérant à des erreurs de quantification. La norme cumulative est une façon d'éviter les erreurs de quantification aux bornes des *bins* en considérant la somme cumulée des différences :

$$d_{CL_1}(\mathbf{P}, \mathbf{Q}) = \sum_i |\tilde{P}_i - \tilde{Q}_i| \text{ avec } \tilde{P}_i = \sum_{j \leq i} P_j \text{ et } \tilde{Q}_i = \sum_{j \leq i} Q_j \quad (3.19)$$

3.4. UN CRITÈRE DE VALIDATION PHOTOMÉTRIQUE DES CANDIDATS

Parmi ces méthodes, on trouve les distances basées sur la forme quadratique :

$$d_A(\mathbf{P}, \mathbf{Q}) = \sqrt{(\mathbf{P} - \mathbf{Q})^\top \mathbf{A} (\mathbf{P} - \mathbf{Q})} \quad (3.20)$$

où \mathbf{A} est une matrice en général symétrique et semi-définie positive, pour que la mesure de distance soit une métrique. Pele *et al.* [68] proposent de combiner les propriétés de cette mesure avec la robustesse du χ^2 aux *bins* de grande cardinalité et démontrent que cette famille de distance conserve les propriétés des deux autres.

Une autre distance qui permet la comparaison *cross-bin* est la distance de *Earth-Mover* qui tient son nom de l'analogie entre la comparaison d'histogramme et le nivellement de la terre. La distance correspond à l'effort minimum qu'il faut produire pour changer les couples intensité-cardinalité de \mathbf{P} vers \mathbf{Q} . On peut définir la distance au sol, $d_{i,j}$, qui correspond à la mesure de distance entre les deux *bins* P_i et Q_j . Cette valeur représente la difficulté de déplacer une valeur d'intensité de la classe i de \mathbf{P} vers la classe j de \mathbf{Q} . En général, on choisit une mesure qui est une métrique et qui se formule simplement. On peut prendre par exemple $d_{i,i} = 0$, $d_{i,i+1} = 1$, etc ... Le flux entre les deux distributions noté $F \in \mathbb{M}_{n \times n}$ correspond à la quantité de valeurs d'intensité déplacées entre le *bin* P_i et le *bin* Q_j , il respecte les trois propriétés suivantes :

- $F_{i,j} \geq 0$, on déplace uniquement de \mathbf{P} vers \mathbf{Q} ;
- $\forall_i \sum_j F_{i,j} \leq P_i$ et $\forall_j \sum_i F_{i,j} \leq Q_j$, on ne déplace pas plus que les effectifs disponibles dans les *bins* ;
- $\sum_{i,j} F_{i,j} = \min(\sum_i P_i, \sum_j Q_j)$, le plus d'intensités possibles sont recalées, autrement dit soit \mathbf{P} est vidé, soit \mathbf{Q} est rempli ;
- $F_{opt} = \min_F (\sum_{i,j} d_{i,j} F_{i,j})$, on effectue le moins d'efforts possibles.

Finalement la distance de Earth-Mover s'exprime :

$$d_{EM}(\mathbf{P}, \mathbf{Q}) = \frac{d_{i,j} F_{opt,i,j}}{F_{opt,i,j}} \quad (3.21)$$

Pele *et al.* montrent que cette distance présente les avantages de bien correspondre à la vision humaine et d'être robuste à des variations d'intensité et une mauvaise quantification [67]. Elle est cependant coûteuse en temps d'exécution, car il est nécessaire de résoudre le problème du transport (Problème de Mong-Kantorovich).

La distance de diffusion tient son nom et son origine des équations de diffusion de la température en physique. Si l'on normalise les histogrammes \mathbf{P} et \mathbf{Q} sur un intervalle continu, avec $x \in I$ on peut poser :

$$d(x) = \mathbf{P}(x) - \mathbf{Q}(x)$$

On introduit la température $T(x, t_0)$, qui suit l'équation :

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} \quad (3.22)$$

$$T(x, t_0) = d(x) \quad (3.23)$$

La distance de diffusion s'exprime alors :

$$d_{diff}(\mathbf{P}, \mathbf{Q}) = \int_{t_0}^{t_f} \text{norm}(T(x, t)) dt \quad (3.24)$$

avec une norme donnée ; Ling *et al.* [47] suggèrent d'utiliser la norme L_1 . t_0 et t_f correspondent à une mesure assimilable au temps, elles sont fixées à l'avance. De même que

pour la distance de *Earth-mover*, cette distance présente l'avantage de bien modéliser le caractère cross-bin mais représente un temps de calcul considérable, car il est nécessaire de calculer cette intégrale itérativement.

3.4.3 Recherche d'un critère invariant en rotation

Dans le cas de notre marqueur circulaire, pour pouvoir directement comparer les images après rectification il est nécessaire de retrouver la rotation afin d'avoir la rectification complète. Le marqueur *bleam* dispose d'un détrompeur entre les deux cercles concentriques, mais nous avons fait le choix de ne pas l'utiliser pour la rectification métrique. On cherche en effet à travailler dans le cadre plus général d'un marqueur à cercles concentriques et en général les marqueurs circulaires ne présentent pas de motifs permettant de résoudre l'ambiguïté de rotation.

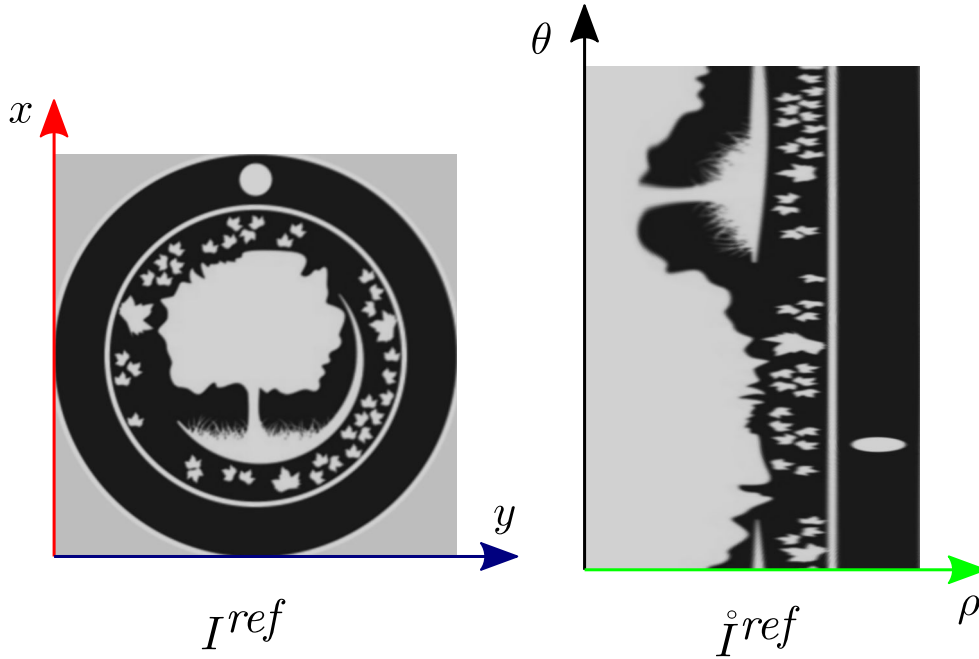


FIGURE 3.3 – Représentation d'une image rectifiée de marqueur en polaire

Nous nous sommes donc intéressés à des mesures invariantes à la rotation du marqueur autour de son centre. Afin de simplifier ce problème d'invariance, nous avons transformé le système de coordonnées cartésiennes naturel de l'image en un système de coordonnées polaires. Ainsi la dimension inconnue, mesurée par l'angle de rotation, θ , de notre problème peut être séparée de l'autre dimension spatiale, mesurée par la distance radiale, ρ . On note l'image en coordonnées polaires de l'image I , l'image \dot{I} avec la correspondance :

$$\dot{I}_{\rho,\theta} = I_{\rho \cos \theta, \rho \sin \theta} \quad (3.25)$$

Cette transformation nécessite une interpolation des images, nous utilisons une interpolation bicubique. La FIGURE 3.3 illustre cette transformation, on peut observer qu'une translation verticale de l'image polaire correspond à une rotation.

3.4. UN CRITÈRE DE VALIDATION PHOTOMÉTRIQUE DES CANDIDATS

La transformation de l'image liée à la rotation autour de l'axe du cercle devient alors une translation dans l'image polaire. Les méthodes de type corrélation de phase ont été utilisées à ces fins pour l'identification et la reconnaissance de l'iris des yeux. En effet la formulation de ce problème est très similaire à la nôtre, chercher à identifier un iris revient en effet à rechercher l'image d'une forme circulaire délimitée par deux cercles, la sclère blanche entourant l'iris et la pupille noire délimitant l'intérieur. Miyazawa *et al.* [52] proposent d'utiliser la mesure de la corrélation de phase pour calculer la similitude entre deux images d'iris normalisées en coordonnées polaires. La corrélation de phase consiste à calculer :

$$\mathcal{T}_R(j, k) = \frac{\mathcal{T}_{\hat{I}_{ref}}(j, k) \overline{\mathcal{T}_{\hat{I}_{rect}}(j, k)}}{|\mathcal{T}_{\hat{I}_{ref}}(j, k) \overline{\mathcal{T}_{\hat{I}_{rect}}(j, k)}|} \quad (3.26)$$

où $\bar{}$ correspond au conjugué et \mathcal{T} représente la transformée de Fourier bidimensionnelle discrète de l'image. On peut ensuite recalculer une image avec la transformée de Fourier inverse, \mathcal{T}^{-1} , on obtient une nouvelle image. L'image obtenue a ainsi la propriété de présenter un pic d'intensité très forte, en (ρ_0, θ_0) , si les deux images peuvent être alignées par une translation de (ρ_0, θ_0) . Cette propriété est très intéressante, car elle permet, en plus de fournir une mesure de l'alignement, de retrouver directement la rotation de l'image en étudiant la valeur de θ_0 . Ces classes de méthodes sont cependant très sensibles au flou de focalisation ou à la différence de résolution [52]. Dans le cas de nos marqueurs où cette mesure servirait à être utilisée pour estimer l'homographie de marqueur pouvant être placé à des distances différentes de la caméra, on constate que cette méthode fonctionne moins efficacement.

Si on mesure les intensités sur chaque colonne de l'image, on peut extraire des informations statistiques qui sont invariantes à une translation sur les lignes de l'image. On peut ainsi travailler sur des mesures invariantes à la rotation du marqueur autour de son centre. Une première mesure de ce type consiste à calculer la moyenne sur chaque colonne de l'image, cette mesure que nous appelons « profil radial » de l'image est représentée sur la figure 3.4.

Il est alors possible de comparer le « profil radial » de l'image du marqueur de référence, \hat{I}^{ref} avec celui de l'image rectifiée, \hat{I}^{rect} . Le profil radial peut alors être traité de la même façon qu'une image en étant considéré comme une valeur unidimensionnelle, on peut mesurer la corrélation de Pearson ou la norme L_2 . À noter qu'il est également possible d'utiliser la médiane à la place de la moyenne.

Il est également possible d'utiliser la distribution des intensités dans chaque colonne, le calcul d'un histogramme contient en effet plus d'information que le profil radial. Soit ρ fixé, correspondant à une colonne des images polaires à comparer. On notera \mathbf{P} et \mathbf{Q} , les vecteurs des *bins* de l'histogramme des deux colonnes à comparer pour cette valeur de ρ . En réalité, on devrait les noter \mathbf{P}_ρ et \mathbf{Q}_ρ , mais pour simplifier les notations dans la suite, nous omettons le suffixe ρ .

$$P_i = \text{card} \left\{ \hat{I}_{\rho, \theta} \text{ tel que } \forall \theta \in [0, 2\pi] \frac{i-1}{n} \leq \frac{\hat{I}_{\rho, \theta}}{255} \leq \frac{i}{n} \right\} \quad (3.27)$$

3.4.4 Création de critères photométriques pour notre méthode

Dans les sections précédentes, nous avons d'abord présenté des critères fondés sur la mesure des intensités et de leur distribution statistique puis nous avons fourni des éléments pour construire un critère photométrique fondé sur l'invariance en rotation. Dans cette section nous allons présenter trois critères photométriques retenus que nous

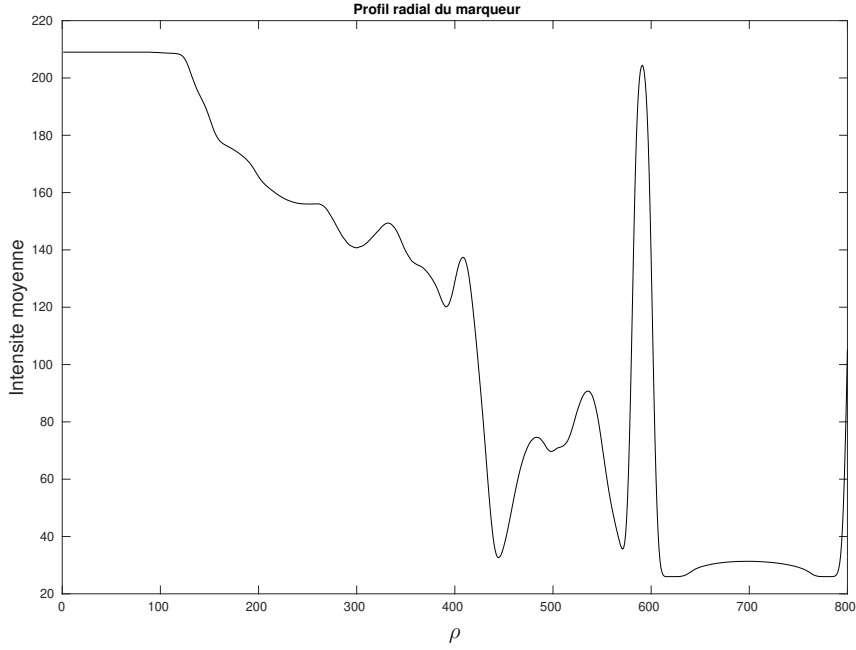


FIGURE 3.4 – Profil radial du marqueur de la FIGURE 3.3

avons utilisés dans notre méthode. Nous avons cherché à identifier parmi les critères de l'état de l'art ceux présentant des caractéristiques pertinentes. Parmi les critères que nous avons retenus et dont l'évaluation est présentée dans la section 3.5 nous avons la corrélation croisée du profil radial, la mesure du χ^2 de l'histogramme radial et la distance de diffusion de l'histogramme radial.

Le premier critère est ce que nous appelons la corrélation croisée du profil radial. Le profil radial, \mathbf{G} , se définit à partir de l'image en coordonnées polaires :

$$G_\rho = \int_{\theta=0}^{2\pi} \dot{I}_{\rho,\theta} \quad (3.28)$$

Pour son calcul dans une image réelle, il faut cependant estimer ce profil de façon discrète. On définit donc deux variables n_r et n_w qui correspondent respectivement à la résolution sur le rayon ρ et sur l'angle θ de l'image polaire. Ainsi nous avons le profil radial discret qui correspond à :

$$G_i = \sum_{j=1}^{n_w} \dot{I}_{(\frac{i}{n_r}\rho_{max}, \frac{j}{n_w}2\pi)} \text{ pour } i \in \llbracket 1, n_r \rrbracket \quad (3.29)$$

Nous pouvons ainsi définir une mesure de la corrélation croisée du profil radial entre l'image de référence du marqueur, I_{ref} , et l'image rectifiée, I_{rect} qui vaut :

$$d_{prcc}(I^{ref}, I^{rect}) = \frac{\sum_{i=1}^{n_r} (G_i^{ref} - \bar{G}^{ref})(G_i^{rect} - \bar{G}^{rect})}{\sqrt{\sum_{i=1}^{n_r} (G_i^{ref} - \bar{G}^{ref})^2 \sum_{i=1}^{n_r} (G_i^{rect} - \bar{G}^{rect})^2}} \quad (3.30)$$

L'intérêt de ce critère est qu'il présente de la robustesse aux changements d'intensité liés à l'éclairage puisqu'il est invariant à une transformation affine des valeurs d'intensités des deux images comparées. Cependant par le calcul de la moyenne, de l'information photométrique est perdue lors de la comparaison.

De façon similaire au profil radial, on peut, au lieu de calculer la moyenne, conserver les informations statistiques au sein d'une même colonne de l'image polaire, puisque celle-ci est invariante à la translation (ou à la rotation de l'image cartésienne). L'histogramme radial, \mathbf{P}_i , peut être également défini de façon discrète avec :

$$P_{i,k} = \text{card} \left\{ \hat{I}_{(\frac{i}{n_r} \rho_{max}, \frac{j}{n_w} 2\pi)} \text{ tel que } j \in \llbracket 1, n_w \rrbracket, \frac{k-1}{n} \leq \frac{\hat{I}_{(\frac{i}{n_r} \rho_{max}, \frac{j}{n_w} 2\pi)}}{255} \leq \frac{k}{n} \right\} \quad (3.31)$$

où k est l'indice associé au *bin* et n le nombre de *bin* choisi pour l'histogramme.

Nous définissons alors une mesure de comparaison des histogrammes radiaux fondée sur le χ^2 :

$$d_{hrchi2} = \sum_{i=0}^{n_r} d_{\chi^2}(\mathbf{P}_i^{ref}, \mathbf{P}_i^{rect}) \quad (3.32)$$

Ce critère permet de prendre en compte plus d'information photométrique sur le marqueur tout en gardant les propriétés d'invariance en rotation que nous recherchons, il n'est cependant pas robuste au changement d'éclairage.

Le dernier critère proposé permet de prendre en compte le changement d'éclairage, car il s'agit d'une mesure *cross-bin*. Nous proposons ainsi la mesure de dissimilitude de l'histogramme radial fondée sur la diffusion :

$$d_{hrdiff} = \sum_{i=0}^{n_r} d_{diff}(\mathbf{P}_i^{ref}, \mathbf{P}_i^{rect}) \quad (3.33)$$

3.5 Une évaluation de notre méthode

Dans cette section, nous proposons une évaluation de notre procédé de génération de candidats pour la rectification métrique. L'idée sous-jacente est de valider en partie notre approche en confirmant que parmi les candidats générés pour la rectification, il est possible d'en trouver un qui rectifie correctement l'image. Pour cela des images d'un marqueur inscrit dans un damier ont été prises par différents appareils que nous avons calibrés de façon précise et auxquelles nous fournissons un calibrage « par défaut » empirique. La qualité de la reconstruction est évaluée à partir de l'erreur de reprojection sur les points d'un damier dans lequel notre marqueur a été inséré. Nous comparerons ainsi l'erreur de reprojection des méthodes directes avec l'erreur minimale obtenue sur les rectifications candidates parmi tous les candidats générés par notre modèle. On procède de la même façon que dans le chapitre 2, la librairie OpenCV est utilisée pour la détection des coins du damier et le calibrage des caméras.

3.5.1 Construction d'un calibrage par défaut

Il n'existe pas de base de données complète qui regroupe toutes les données des paramètres intrinsèques « usine » des smartphones. Le site <https://www.digicamdb.com> présente une base de données des appareils photographiques mais celle-ci ne concerne pas les téléphones mobiles.

Nous sommes partis du constat que pour les smartphones actuels, la taille du capteur en pixel/mm et la focale de la lentille sont souvent similaires d'un appareil à un autre. De ce fait, il est possible de constituer un calibrage « par défaut » avec les appareils que nous avons utilisés. Nous définissons une fonction empirique s'appuyant sur la taille de la diagonale de l'image. Si nous calculons la valeur $\frac{f_x}{\sqrt{n^2+m^2}}$ où n et m sont la résolution en hauteur et largeur de l'image, nous constatons qu'elle varie peu selon les appareils. Cette valeur est également proportionnelle à la distance focale équivalente en 35mm, une métrique utilisée dans le milieu de la photographie.

$$f_{35} = f_x \frac{\sqrt{24^2 + 36^2}}{\sqrt{n^2 + m^2}} \quad (3.34)$$

Après avoir calibré une douzaine d'appareils photo et obtenu les données des fabricants de vingt autres appareils photo, nous estimons un modèle gaussien de l'équivalent 35mm de la distance focale, tel qu'illustré dans la FIGURE 3.5. Plus précisément, nous estimons une fonction gaussienne (en bleu) basée sur les valeurs collectées ou estimées (en rouge) sur la focale de différents modèles de *smartphones*. Dans notre cas, nous avons obtenu

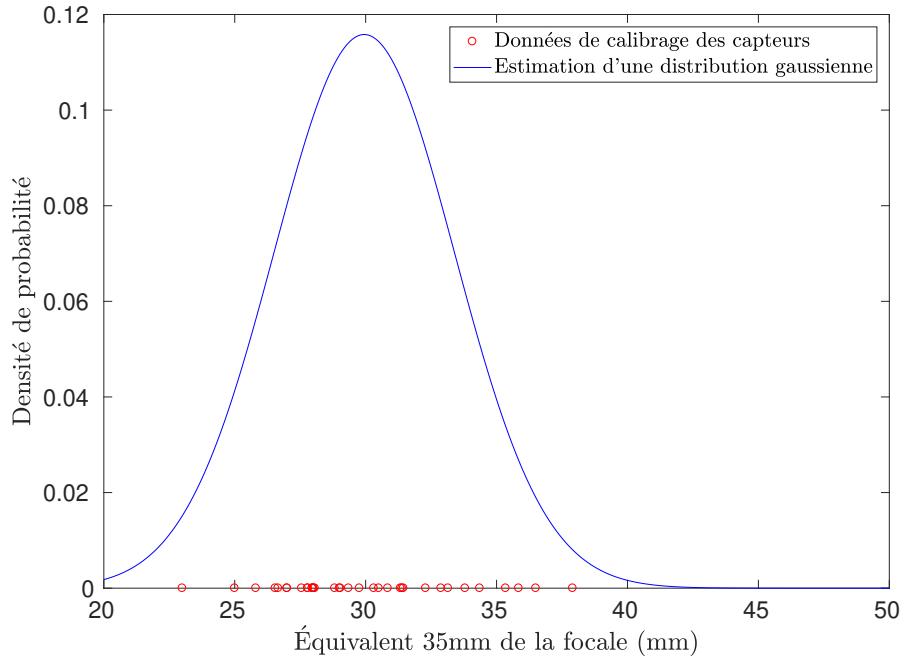


FIGURE 3.5 – Calcul d'un modèle par défaut à partir du calibrage de plusieurs caméras

expérimentalement une focale moyenne de $f_{35} = 2,995 \times 10^1$ mm avec une variance de $\sigma_{f_{35}}^2 = 1,186 \times 10^1$. On fait de même pour l'image du point principal en normalisant par la résolution de l'image.

$$u_m = \frac{u_0 - \frac{m}{2}}{m} \quad (3.35)$$

$$v_n = \frac{v_0 - \frac{n}{2}}{n} \quad (3.36)$$

3.5. UNE ÉVALUATION DE NOTRE MÉTHODE

Algorithme	Paramètres de calibrage	Critère de sélection des candidats
Vérité terrain	calibrés	n/a
<i>EK</i> calibré	calibrés	n/a
<i>EX</i> attendu	par défaut + incertitudes	erreur de reprojection du damier
<i>EX</i> prcc	par défaut + incertitudes	corrélacion croisée des profils radiaux
<i>EX</i> hrchi2	par défaut + incertitudes	mesure du χ^2 sur les histogrammes radiaux
<i>EX</i> hrdiff	par défaut + incertitudes	distance de diffusion sur les histogrammes radiaux
<i>EK</i> défaut	par défaut	n/a

TABLE 3.1 – Différents algorithmes testés et comparés

Ce qui permet également de calculer une distribution gaussienne et d'obtenir l'incertitude de ces paramètres avec plusieurs caméras. On obtient $u_m^- = -6,510 \times 10^{-3}$, $\sigma_{u_m}^2 = 1,333 \times 10^{-3}$ et $v_m^- = -2,220 \times 10^{-4}$ et $\sigma_{v_m}^2 = 2,466 \times 10^{-4}$.

3.5.2 Évaluation des candidats générés

Méthodes comparées

Dans cette section nous évaluons les méthodes du TABLEAU 3.1. Ces méthodes sont classées selon le degré de connaissance sur les paramètres et les ellipses utilisés. Les méthodes qu'on note *EX* font référence à notre approche qui s'appuie sur le calibrage « par défaut » et la détection d'une seule ellipse afin de déterminer un espace de candidat réduit. La différence entre les méthodes *EX* du tableau tient dans le critère de sélection des rectifications candidates générées à partir des candidats pour l'image du centre. La méthode « *EX* attendu » est un peu particulière, son critère utilise en fait l'erreur de reprojection des points du damier. Ce que cette méthode vérifie, c'est qu'il existe bien dans l'ensemble des $N = 200$ candidats générés, un candidat qui rectifie correctement l'image. Les autres méthodes reprennent celles vues dans le chapitre 2.

Mesure de l'erreur

Pour mesurer l'erreur nous utilisons la même mesure que dans le chapitre 2, c'est-à-dire l'erreur de reprojection sur les points du damier. On présente sur la FIGURE 3.6, la distribution des erreurs mesurées sur nos méthodes.

Commentaires sur les résultats

Le premier constat est que la méthode "*EX* attendu" atteint la même précision que lorsque la caméra est calibrée, ce qui tend à confirmer notre modèle de distribution proposé pour l'image du centre. De plus s'il est possible de déterminer une image du centre qui rectifie mieux, c'est qu'avec un critère idéal, il est possible de déterminer la rectification idéale avec l'algorithme que nous avons proposé. Cependant, nous faisons également un constat plus négatif, car aucun des critères que nous avons testés ne l'atteint, au contraire, les rectifications choisies donnent de très grandes erreurs de reprojection. Pour illustrer ce problème, nous avons affiché les rectifications sur la FIGURE 3.7. On peut observer qu'il y a peu de différence entre la rectification calculée en maximisant la corrélation des profils radiaux (cf. FIGURE 3.7b) et la rectification calculée à partir de la vérité terrain (cf. FIGURE 3.7a) malgré la différence en erreur de projection sur le damier bien plus importante. De plus, il faut rappeler que l'invariance en rotation nous

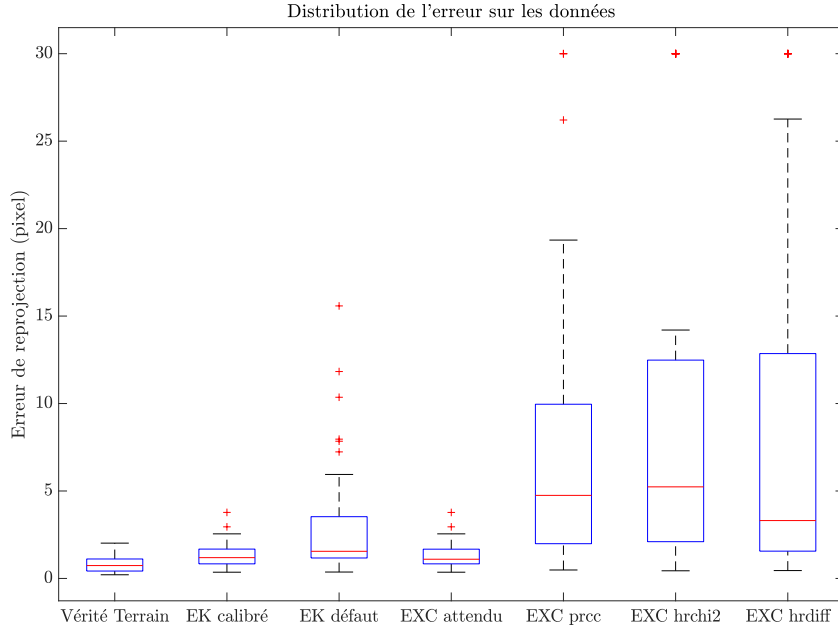


FIGURE 3.6 – Erreur de projection sur les méthodes testées

impose des critères qui ne peuvent prendre en compte que les informations statistiques sur chaque colonne.

Ces résultats montrent qu'il n'existe pas de critère photométrique suffisant pour résoudre notre problème. Nous pensons que cela signifie qu'il faut rejeter au moins l'une de nos deux hypothèses :

- Un critère fondé sur l'invariance en rotation est suffisamment discriminant pour localiser l'image du centre du marqueur.
- L'intérieur du marqueur, lorsqu'il est visible, est suffisamment discriminant pour localiser précisément son centre.

Lorsque le marqueur se situe à une distance importante de la caméra, la perspective s'approche d'une transformation affine et l'image du centre du marqueur s'approche du centre de l'image du contour. Ainsi, même l'erreur de reprojection de points à l'intérieur du marqueur ne contraint plus suffisamment la perspective pour l'estimer convenablement.

Pour répondre à ces problèmes, voici les pistes proposées et qui seront explorées dans le chapitre 5 :

- soit estimer, dans un premier temps la rotation, afin d'éviter le problème d'invariance en rotation ;
- soit utiliser un critère photométrique, mais en exploitant des données extérieures aux marqueurs.

3.6 Conclusion et perspectives sur la méthode proposée

Nous avons proposé une méthode de rectification du plan de support d'un marqueur circulaire qui ne nécessite la détection que d'une seule ellipse lorsqu'un modèle de cali-

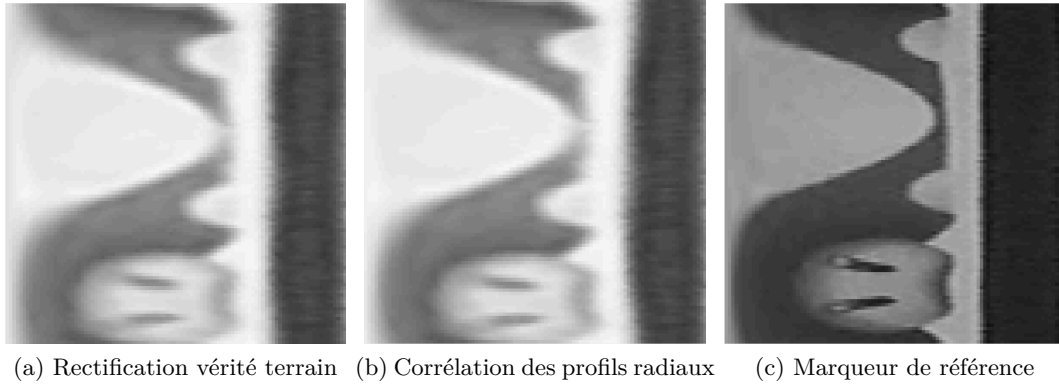


FIGURE 3.7 – Exemple de la rectification d'un marqueur

brage par défaut de l'appareil peut être construit. L'idée de l'approche consiste à utiliser des *a priori* sur les paramètres de la caméra et à construire une estimation des distributions possibles de l'image du centre du marqueur. On rappelle que l'image du centre est en relation polaire avec la ligne de fuite par rapport à l'image du cercle contour. Ainsi, il est possible de rechercher pour chaque position probable de l'image du centre des rectifications métriques que nous pouvons ensuite évaluer. Nous avons ensuite étudié plusieurs critères photométriques pour évaluer ces solutions. Si nous n'avons malheureusement pas réussi à mettre en avant un critère efficace pour sélectionner les candidats, nous avons mis en avant que notre méthode de proposition de candidats permet de proposer des rectifications métriques précises.

Dans les prochains chapitres, nous proposons de poursuivre notre approche en proposant, dans le chapitre 4, de travailler sur la double ambiguïté dans le calcul de la normale du plan de support. Dans le chapitre 5, nous étudierons comment il est possible d'utiliser de l'information à l'extérieur du marqueur pour estimer la rotation autour de l'axe du cercle afin de contourner les dernières difficultés rencontrées au sein de ce chapitre.

Chapitre 4

Étude de l'ambiguïté de la pose d'un cercle vu par une caméra

Sommaire

4.1	Ambiguïté de la pose du point de vue d'une caméra monoculaire	66
4.2	Nouvelle formulation du calcul de la pose d'un cercle . . .	68
4.2.1	Calcul de la ligne de fuite à partir des coniques dégénérées . . .	69
4.2.2	Une observation intéressante sur les points de base	73
4.3	Paramétrage minimal pour étudier l'ambiguïté	74
4.3.1	Description du paramétrage	74
4.3.2	Une première condition nécessaire et suffisante	75
4.3.3	Une condition suffisante exploitable	76
4.4	Conclusion sur l'étude de l'ambiguïté	77

4.1 Ambiguïté de la pose du point de vue d'une caméra monoculaire

Dans les chapitres précédents, nous avons montré qu'il était possible de retrouver la structure euclidienne d'un plan de support à partir de l'image d'un cercle et des paramètres intrinsèques de la caméra. Cependant, la résolution du problème ne donne pas une solution unique mais deux solutions algébriques, ce qui introduit une ambiguïté. Il est possible de supprimer cette ambiguïté en utilisant un modèle d'apparence photométrique du marqueur, comme vu précédemment, en s'inspirant des méthodes proposées par Bergamasco et Pagani qui valident une des deux poses ambiguës en se basant sur leur propre critère photométrique [4, 63]. Cependant, dans ce chapitre, nous souhaitons aborder une étude plus théorique de cette ambiguïté. Nous rappelons que cette ambiguïté apparaît dans le calcul de la pose d'un cercle à partir de sa projection perspective. Il est évident que l'ambiguïté ne peut être résolue sans information supplémentaire. Ainsi, ce que nous allons voir dans ce chapitre c'est comment formuler une condition minimale sur les paramètres de la pose du cercle permettant de savoir dans certains cas quelle solution algébrique correspond à la ligne de fuite.

Dans le cas de l'image d'un cercle seul, il est connu qu'il existe une double solution pour la normale du plan de support (et donc pour la pose) si la caméra est calibrée [81]. Dans le cadre de ce chapitre, nous utilisons le mot pose pour désigner la pose d'un cercle (le cercle externe dans le cas du marqueur), le problème traité est ainsi plus généralisable à n'importe quelle estimation de la pose d'un cercle. À noter que celle-ci n'est pas équivalente à la pose d'une caméra puisque celle-ci requiert une rotation supplémentaire autour de l'axe du cercle. Il s'agit donc d'un problème à 5 degrés de liberté : 3 pour la translation et seulement 2 pour l'orientation du cercle.

La contribution est de fournir un nouveau cadre géométrique pour énoncer le problème posé, à savoir comment calculer la pose du cercle à partir de son image et des paramètres intrinsèques de la caméra. Nous étudierons en particulier la double ambiguïté de la pose du cercle dans ce problème. Le premier objectif est de donner les solutions au problème de la pose du cercle à partir de son image et nous démontrons les résultats théoriques présentés au chapitre 3. Le second objectif est de fournir une condition paramétrée par un ensemble minimal de paramètres sur la pose¹ qui permette de lever l'ambiguïté. Pour cela, nous remarquons que lorsque l'on est capable de déterminer si 2 points particuliers du plan de support, que nous nommons les points de base, sont placés tous deux du même côté de la caméra (condition quasi-affine) alors nous sommes capables de distinguer la solution algébrique correcte parmi les deux solutions ambiguës. Nous présentons en fin de chapitre un exemple de configuration remarquable pour illustrer les principes introduits.

Projection d'un cercle par une caméra perspective et affine

On se propose de rappeler les fondements géométriques du problème de l'estimation de la pose d'un cercle et d'introduire les notations utilisées dans ce chapitre.

Soit \mathcal{P} , le plan de support du cercle, d'équation $z = 0$ dans le repère monde.

Nous considérons la matrice de projection d'une caméra perspective : $P \sim R [I \mid \mathbf{T}]$, où la matrice de rotation $R \in \mathbb{SO}_3$ et le vecteur de translation $\mathbf{T} \in \mathbb{R}^3$ décrivent la pose de la caméra, c'est-à-dire respectivement son orientation et la position dans repère monde. À noter qu'on omet la matrice des paramètres intrinsèques, de façon générale celle-ci

1. Évidemment, si tous les paramètres de pose sont connus, l'ambiguïté est toujours levée.

sera omise sans perte de généralité dans tout ce chapitre, car on se place uniquement dans le cas calibré.

La pose de \mathcal{P} dans le repère de la caméra est donnée par le vecteur $\mathbf{N} = [\mathbf{r}_3^\top, -d]^\top$, où \mathbf{r}_3 est la troisième colonne de \mathbf{R} qui définit la normale du plan \mathcal{P} , et d est la distance orthogonale du centre optique de la caméra au plan \mathcal{P} . La restriction à \mathcal{P} de la projection est une homographie dont la matrice s'écrit $\mathbf{H} \sim \mathbf{R} [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \mathbf{T}]$ où \mathbf{e}_1 et \mathbf{e}_2 sont les deux premières colonnes de la matrice identité d'ordre 3. Dans le plan projectif, n'importe quelle conique peut être représentée en coordonnées homogènes 2D par une matrice symétrique réelle d'ordre 3. On rappelle que pour une projection perspective, tout cercle de \mathcal{P} , en supposant l'hypothèse d'invariance quasi-affine [30, p. 515], c'est-à-dire que tous ses points se trouvent devant la caméra, est transformé par l'homographie \mathbf{H} en une ellipse par l'équation de projection $\mathbf{C} = \mathbf{H}^{-\top} \mathbf{Q} \mathbf{H}^{-1}$, où $\mathbf{Q} \in \mathcal{S}_3(\mathbb{R})$ ² est la matrice du cercle et $\mathbf{C} \in \mathcal{S}_3(\mathbb{R})$ est la matrice de l'ellipse.

Supposons maintenant le modèle simplifié d'une caméra affine ou *weak perspective* [30, P. 171]. Ce modèle est intéressant car il fournit une bonne approximation dans le cas où le cercle est observé à une distance importante de la caméra, mais surtout parce qu'il permet de visualiser intuitivement l'ambiguïté de la pose du cercle.

La matrice de projection par une caméra affine est alors :

$$\mathbf{P}_a \sim \begin{bmatrix} \mathbf{R}^{1\top} & \mathbf{T}_1 \\ \mathbf{R}^{2\top} & \mathbf{T}_2 \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (4.1)$$

Dans le cas calibré, sans perte de généralité, cette matrice est celle d'une projection orthographique.

Origine de l'ambiguïté de pose du carré et du cercle

Dans le cadre d'un contour carré à l'instar des marqueurs ARTag [19], l'image par une caméra affine est un parallélogramme et il existe alors une ambiguïté dans l'estimation de la pose du carré (en plus de l'ambiguïté liée à sa symétrie). La FIGURE 4.1 présente un exemple de pose ambiguë, à partir de la même projection sur le plan image \mathcal{I} où deux poses sont possibles.

Cette ambiguïté n'existe pas dans le cas d'une projection perspective lorsque l'image d'un carré est un quadrilatère. Cependant lorsque la distance entre l'objet et la caméra est importante ou à mesure que la distance focale grandit, la projection perspective tend vers une projection affine. C'est d'ailleurs, l'hypothèse faite par Köser *et al.*, pour calculer la perspective, à partir de la donnée d'une transformation affine locale (généralement apportée par un descripteur). Ainsi les auteurs proposent de considérer la transformation affine locale comme une approximation de la jacobienne d'une transformation perspective, l'identification des coefficients fournit alors la pose du plan [44]. On peut ainsi constater que cette méthode fournit jusqu'à 4 solutions. Ces solutions ambiguës représentent donc des solutions perspectives proches de la transformation affine observée. Une autre façon de le dire est que lorsque l'homographie est proche d'une transformation affine, au sens de l'erreur de reprojection, les poses ambiguës de la résolution affine donnent des solutions perspectives proches de la transformation observée.

Ainsi certains auteurs constatent que lorsque la mesure de la perspective s'avère difficile par exemple à cause d'un manque de précision de mesure sur l'image, une ambiguïté

2. $\mathcal{S}_3(\mathbb{R})$ est l'ensemble des matrices symétriques réelles d'ordre 3

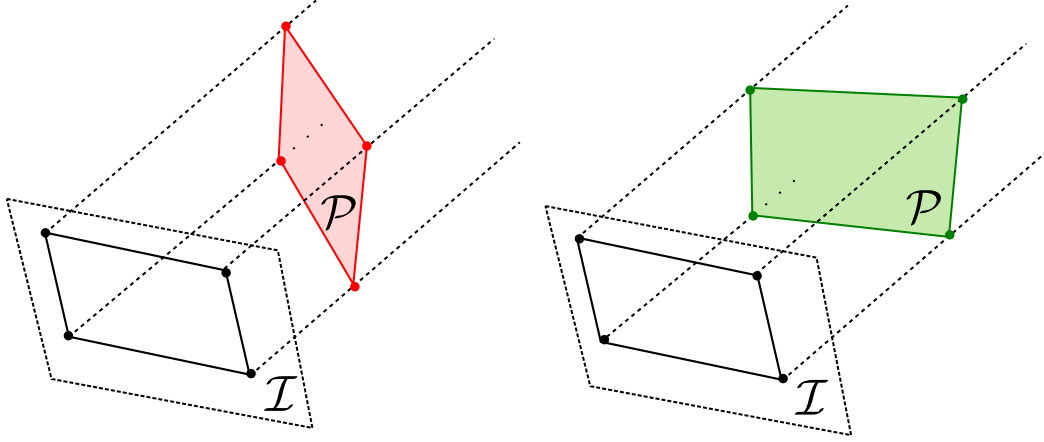


FIGURE 4.1 – Représentation de l'ambiguïté sur la pose du carré avec une caméra affine

peut alors apparaître [79, 72]. Schweighofer *et al.* montrent qu'il est possible de calculer les deux positions ambiguës correspondantes et proposent ainsi d'évaluer l'erreur de reprojection dans les deux cas. Ils montrent ainsi que l'erreur de reprojection admet parfois deux minima locaux, qu'il convient de prendre en compte pour une estimation robuste.

Concernant la projection du cercle, l'ambiguïté existe à la fois dans le cas affine et le cas perspectif. À partir de la FIGURE 4.2, on peut se donner l'intuition qu'il existe deux façons de positionner le cercle sur le cylindre elliptique (rétro-projection de l'image du cercle) dans le cas affine, ou sur le cône elliptique dans le cas perspectif. Nous verrons plus tard dans ce chapitre la configuration des poses ambiguës pour le cas projectif. Dans le cas affine, les deux poses du cercle sont en fait symétriques par rapport au plan perpendiculaire au plan image et passant par le grand axe de l'ellipse. L'observation géométrique de l'ambiguïté dans le cas perspectif a été constatée, mais celle-ci a finalement été peu étudiée en détail. Zheng *et al.* [84] mentionnent l'existence des deux solutions algébriques, correspondant aux solutions géométriques indiscernables, mais sans fournir une compréhension claire de la façon dont l'ambiguïté peut être levée. Pagani *et al.* [63] présentent une méthode assez similaire à la solution proposée par Chen *et al.* [10], où l'ambiguïté est résolue en minimisant la distance entre l'image du marqueur rectifié et le motif attendu sur toutes les poses possibles.

4.2 Nouvelle formulation du calcul de la pose d'un cercle

Dans un premier temps, reprenons l'estimation de la pose d'un cercle, ce problème peut être traité par la séquence suivante :

- calculer l'homographie, paramétrée par l'image du cercle et la ligne de fuite,
- extraire la pose du cercle à partir de l'homographie.

Comme vu dans le chapitre 2, on sait qu'une homographie H peut être déterminée à partir de la connaissance de l'image du cercle C et de la ligne de fuite \mathbf{v}_∞ du plan de support, \mathcal{P} . Soit $S_{\mathcal{P}}$, une similarité du plan de support, \mathcal{P} , qui transforme le cercle Q en un cercle unitaire centré sur l'origine et $S_{\mathcal{I}}$ une similarité du plan image \mathcal{I} qui transforme la matrice C sous une forme canonique telle qu'on a : $C' = \text{diag}(C'_{11}, C'_{22}, C'_{33})$. En utilisant une approche similaire à [7] avec la notation $[u, v, 1]^\top \sim C^{-1}\mathbf{v}_\infty$, il peut être

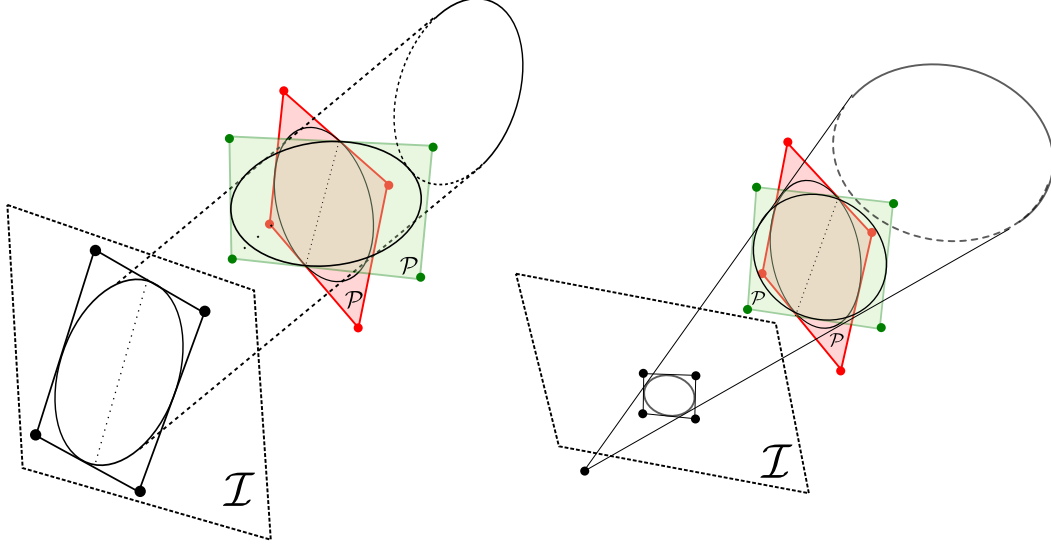


FIGURE 4.2 – Représentation de l’ambiguïté sur la pose du cercle avec une caméra affine et perspective

démontré que, dans l’hypothèse d’une caméra avec des pixels carrés, on a $\mathbf{H} \sim \mathbf{S}_{\mathcal{I}}^{-1} \mathbf{M} \mathbf{S}_{\mathcal{P}}$ où

$$\mathbf{M} = \begin{bmatrix} -1 & C'_{22}uv & -u \\ 0 & -C'_{11}u^2 + 1 & -v \\ -C'_{11}u & C'_{22}v & 1 \end{bmatrix} \begin{bmatrix} r & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & s \end{bmatrix}$$

$$\text{avec } r = \left(-\frac{C'_{22}}{C'_{11}} C'_{11}(C'_{11}u^2 + C'_{22}v^2 + C'_{33})\right)^{1/2} \text{ and } s = \left(-C'_{22}(1 - C'_{11}u^2)\right)^{1/2}$$

Notez que les matrices $\mathbf{S}_{\mathcal{P}}$ et $\mathbf{S}_{\mathcal{I}}$ peuvent être entièrement déterminées par l’image du cercle, \mathbf{C} , \mathbf{M} peut l’être également à une rotation 2D près (autour du centre du cercle sur \mathcal{P}).

Sous cette forme il apparait clair que les difficultés du problème et notamment l’ambiguïté sur la pose proviennent en fait de la résolution du calcul de la ligne de fuite du plan qui nous donne deux solutions. Nous proposons donc de revenir en détail sur la méthode de la section 3.2 présentée dans le chapitre précédent pour mieux comprendre cette ambiguïté. Ensuite, nous montrerons à l’aide d’un nouveau cadre géométrique que dans certaines conditions, nous pouvons déterminer des configurations géométriques simples où l’ambiguïté est résolue.

4.2.1 Calcul de la ligne de fuite à partir des coniques dégénérées

Pour calculer la ligne de fuite, nous reprenons l’idée qu’à partir de deux coniques, il est possible de définir un faisceau dont les membres dégénérés possèdent tous les points d’intersection des deux coniques, et donc les points cycliques pour le cas des cercles.

Soit \mathbf{Q} la matrice d’un cercle sur le plan, \mathcal{P} , et $\psi = \mathbf{H}^\top \omega \mathbf{H}$, celle de la rétroprojection sur \mathcal{P} de l’image de la conique absolue [30, p. 81] donnée par $\omega = \mathbf{K}^{-\top} \mathbf{K}^{-1}$. La conique absolue est une conique virtuelle³ —avec signature $(3, 0)$ — dans le plan à l’infini. Tout

3. Les coniques virtuelles ont pour matrices des matrices positives et définies, elles ne contiennent donc pas de points réels.

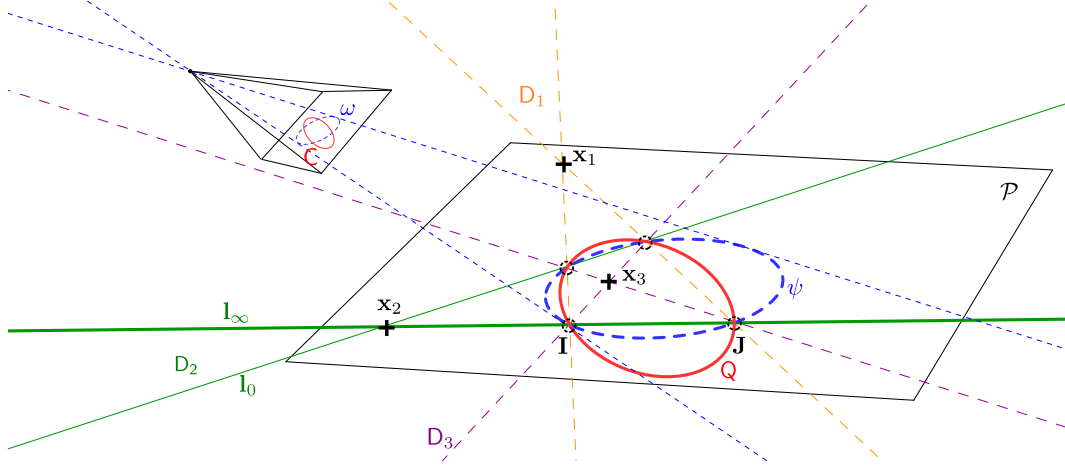


FIGURE 4.3 – Représentation schématique des coniques dégénérées dans le plan \mathcal{P} .

plan affine 3D (c'est-à-dire tout plan projectif autre que le plan à l'infini) intersecte la conique absolue aux points cycliques.

Proposition 4.2.1. *La conique virtuelle ψ obtenue par la rétroprojection de l'image ω de la conique absolue sur n'importe quel plan (affine) est un cercle de ce plan, dans le sens où ψ inclut les points cycliques.*

Démonstration. La preuve est assez simple puisque la conique absolue est le lieu des points cycliques de tous les plans affines 3D et sa rétroprojection sur un plan contient les points cycliques de ce plan, c'est un cercle par définition. \square

Il est à noter que la matrice ψ ne dépend que du vecteur \mathbf{T} , puisque l'on a :

$$\psi = \mathbf{H}^\top \omega \mathbf{H} = \mathbf{T}^\top \mathbf{T} \text{ où } \mathbf{T} = [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \mathbf{T}] \quad (4.2)$$

Soit $\{\alpha_i\}_{i=1,3}$, l'ensemble des valeurs propres généralisées de la paire de matrices (\mathbf{Q}, ψ) , c'est-à-dire les trois racines de l'équation caractéristique $\det(\mathbf{Q} - \alpha\psi) = 0$. L'ensemble des matrices $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ définit un faisceau conique [27] qui inclut trois coniques dégénérées de matrices $\mathbf{D}_i = \mathbf{Q} - \alpha_i\psi$. Ces matrices de rang 2 représentent des paires de droites et ont la forme suivante $\mathbf{D}_i = \mathbf{l}_a^i (\mathbf{l}_b^i)^\top + \mathbf{l}_b^i (\mathbf{l}_a^i)^\top$, où \mathbf{l}_a^i et \mathbf{l}_b^i sont les vecteurs de ces droites. Ces paires de droites sont représentées dans la FIGURE 4.3. Une telle matrice de paires de droites \mathbf{D}_i peut être facilement décomposée pour obtenir les vecteurs des droites bien qu'il soit impossible de distinguer \mathbf{l}_a^i de \mathbf{l}_b^i .

Proposition 4.2.2. *Les signatures projectives⁴ des trois membres dégénérés du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ sont $(1, 1)$, $(2, 0)$ et $(2, 0)$.*

Avant la démonstration, énonçons ce résultat important :

Remarque 4.2.1. *Soit $(\mathbf{A}_1, \mathbf{A}_2)$ et $(\tilde{\mathbf{A}}_1, \tilde{\mathbf{A}}_2)$ deux paires de matrices réelles carrées et d'ordre 3 étant congruentes par rapport à une matrice inversible $\mathbf{H} \in \mathbb{R}^{3 \times 3}$ par la relation $\tilde{\mathbf{A}}_j = s_j \mathbf{H}^{-\top} \mathbf{A}_j \mathbf{H}^{-1}$, avec un scalaire non nul $s_j \neq 0$, pour $j = 1..2$. On peut admettre*

4. La signature projective d'une conique est $\sigma(\mathbf{C}) = (\max(p, n), \min(p, n))$, où p et n sont respectivement les valeurs propres positives et négatives de sa matrice (réelle) \mathbf{C} . Elle est laissée inchangée par les transformations projectives.

que si $(\tilde{\lambda}, \tilde{\mathbf{z}})$ est une paire propre généralisée de $\{\mathbf{A}_1, \mathbf{A}_2\}$, alors $(s_1/s_2, \tilde{\lambda}\mathbf{H}\tilde{\mathbf{z}})$ est une paire propre généralisée de $\{\tilde{\mathbf{A}}_1, \tilde{\mathbf{A}}_2\}$. Il s'ensuit que les vecteurs propres associés aux valeurs propres $\tilde{\lambda}$ sont projectivement invariants à un facteur d'échelle et une permutation près.

Démonstration. D'une part, cf. rappel 4.2.1, les paramètres des trois membres dégénérés du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ sont les mêmes que les valeurs propres généralisées de la paire de matrices (\mathbf{Q}, ψ) . Ces valeurs propres généralisées sont aussi les mêmes que celles de la paire de coniques $(\mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}, \mathbf{l})$, c'est-à-dire qu'elles sont les valeurs propres ordinaires de $\mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}$. Elles sont donc toutes réelles et distinctes tant que le centre du cercle n'est pas sur la droite principale.

D'autre part, un résultat bien connu en algèbre dit que, si $\{\gamma_j\}_{j=1..3}$ désigne l'ensemble des valeurs propres ordinaires d'une matrice réelle $\mathbf{A} \in \mathcal{S}_3(\mathbb{R})$, alors l'ensemble des valeurs propres ordinaires des matrices de rang 2, $\mathbf{A} - \gamma_j \mathbf{l}$, sont respectivement $\{0, \gamma_2 - \gamma_1, \gamma_3 - \gamma_1\}$, $\{0, \gamma_1 - \gamma_2, \gamma_3 - \gamma_2\}$ et $\{0, \gamma_1 - \gamma_3, \gamma_2 - \gamma_3\}$. Sans perte de généralité, supposons que $\gamma_1 > \gamma_2 > \gamma_3$; on vérifie facilement que les signatures de $\mathbf{A} - \gamma_j \mathbf{l}$ sont respectivement $(2, 0)$, $(1, 1)$ et $(2, 0)$. \square

Proposition 4.2.3. *Supposons, sans perte de généralité, que \mathbf{D}_2 est la conique dégénérée du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ avec signature $(1, 1)$, les deux autres coniques dégénérées \mathbf{D}_1 et \mathbf{D}_3 — de signatures $(2, 0)$ et $(2, 0)$ — sont des paires de deux droites complexes conjuguées.*

Notre premier résultat clé est que \mathbf{D}_2 est une paire de droites réelles distinctes, l'une d'elles étant la droite à l'infini, $\mathbf{l}_\infty = [0, 0, 1]^\top$, l'autre est une droite réelle du plan qu'on notera \mathbf{l}_0 .

Démonstration. Le cercle réel \mathbf{Q} et le cercle virtuel ψ ont quatre points d'intersections complexes, parmi lesquels la paire de points cycliques, (\mathbf{I}, \mathbf{J}) et une autre paire de points complexes et conjugués. Tous les membres du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ — y compris les dégénérées — contiennent ces quatre points. Par conséquent, l'une des coniques dégénérées est formée de deux droites passant respectivement par (\mathbf{I}, \mathbf{J}) pour l'une et par la seconde paire de points, pour l'autre. Puisque les deux paires de points sont complexes et conjugués, les deux droites sont réelles et l'une d'elles est la droite à l'infini $\mathbf{l}_\infty \sim \mathbf{I} \wedge \mathbf{J}$, l'autre étant \mathbf{l}_0 .

Notez qu'en écrivant :

$$\mathbf{l}_\infty = [0, 0, 1]^\top \text{ et } \mathbf{l}_0 = [a, b, c]^\top \text{ avec } a^2 + b^2 \neq 0 \quad (4.3)$$

il est facile de vérifier que si $\mathbf{D}_2 = \mathbf{l}_\infty \mathbf{l}_0^\top + \mathbf{l}_0 \mathbf{l}_\infty^\top = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ a & b & 2c \end{bmatrix}$ alors le produit des valeurs propres non nulles de \mathbf{D}_2 est égal à $-(a^2 + b^2)$. Cela signifie, que sa signature est bien $(1, 1)$: ses valeurs propres sont de signes différents. \square

La FIGURE 4.3 présente une illustration du cercle sur son plan de support et des différentes coniques dégénérées. Les trois points de base \mathbf{x}_i , où les droites de chaque paire s'intersectent, sont réels, on peut les observer sur la FIGURE 4.3. De plus, leurs vecteurs directeurs sont les vecteurs propres généralisés de (\mathbf{Q}, ψ) et satisfont $\mathbf{D}_i \mathbf{x}_i = 0$.

De même, dans le plan image, si \mathbf{C} désigne l'image du cercle \mathbf{Q} , l'ensemble de matrices $\{\mathbf{C} - \lambda\omega\}_{\lambda \in \mathbb{R} \cup \{\infty\}}$ définit également un faisceau de coniques dont les membres sont les images du faisceau $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$. Ainsi, la paire de droites $\{\mathbf{C} - \lambda\omega\}$ qui inclut

l'image de \mathbf{l}_∞ autrement dit la ligne de fuite \mathbf{v}_∞ , peut toujours être identifiée puisque c'est le seul membre dégénéré de signature $(1, 1)$. Néanmoins, à cette étape, il est impossible de distinguer \mathbf{v}_∞ de l'autre ligne \mathbf{v}_o , image de \mathbf{l}_o .

Si on suppose que toutes les matrices \mathbf{Q} , \mathbf{C} , ψ et ω sont normalisées de sorte que leur déterminant soit unitaire; on sait que, dans ce cas, les paramètres du faisceau satisfont $\alpha = \lambda$, donc, les valeurs propres généralisées de la paire de matrices (\mathbf{Q}, ψ) sont exactement les mêmes que ceux de (\mathbf{C}, ω) .

Proposition 4.2.4. *Supposons que les trois valeurs propres généralisées $\{\lambda_j\}_{j=1..3}$ liées au faisceau de coniques $\{\mathbf{C} - \beta\omega\}_{\beta \in \mathbb{R} \cup \{\infty\}}$ soient triées de façon à avoir $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Alors $\mathbf{L}_2 = \mathbf{C} - \lambda_2\omega$ est la matrice dégénérée de rang 2 de signature $(1, 1)$ et, de plus, il existe $\alpha \in [0, \frac{\pi}{3}]$ et $p > 0$ tel que l'on a :*

$$\lambda_1 - \lambda_2 = 3p(\cos \alpha - \frac{\sqrt{3}}{3} \sin \alpha) \quad (4.4)$$

$$\lambda_2 - \lambda_3 = 3p(2\frac{\sqrt{3}}{3} \sin \alpha) \quad (4.5)$$

Démonstration. Il suffit de prouver la proposition pour les trois valeurs propres généralisées du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$, cf. rappel 4.2.1. Comme expliqué précédemment, elles sont les même que celles de la paire de matrices $(\mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}, \mathbf{I})$, c'est-à-dire qu'elles sont les valeurs propres ordinaires de $\mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}$, elles sont donc réelles et distinctes. Si \mathbf{w}_j désigne les vecteurs propres ordinaires de $\mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}$ associés à la j -ième valeur propre ordinaire, alors les points de base réels (intersections des paires de droites) ont pour vecteurs $\mathbf{x}_j = \mathbf{T}^{-1}\mathbf{w}_j$.

Les valeurs propres de $\mathbf{E} = \mathbf{T}^{-\top}\mathbf{Q}\mathbf{T}^{-1}$ peuvent être obtenues directement en suivant la formulation inspirée par Kopp *et al.* [43] et en utilisant la propriété hermitienne de \mathbf{E} :

$$\lambda_1 = 2p \cos \alpha + q \quad (4.6)$$

$$\lambda_2 = p(\sqrt{3} \sin \alpha - \cos \alpha) + q \quad (4.7)$$

$$\lambda_3 = -p(\sqrt{3} \sin \alpha + \cos \alpha) + q \quad (4.8)$$

où

$$\begin{aligned} \alpha &= \frac{1}{3} \arccos \left(\frac{1}{2} \det \frac{1}{p} \mathbf{F} \right) \in \left[0, \frac{\pi}{3} \right] \\ q &= \frac{1}{3} \text{trace}(\mathbf{E}) \\ \mathbf{F} &= \mathbf{E} - q\mathbf{I} \\ p &= \sqrt{\text{trace} \left(\frac{1}{6} (\mathbf{F})^2 \right)} \end{aligned}$$

Puisque la conique \mathbf{E} est définie à un facteur d'échelle non nul près, nous pouvons choisir un facteur tel que $\text{trace}(\mathbf{E})$ soit positif. On montre alors que λ_2 est le paramètre de la conique de signature $(1, 1)$. Comme expliqué plus tôt (voir la preuve de la proposition 4.2.2) l'ensemble des valeurs propres ordinaires de la matrice de rang 2, $\mathbf{E} - \lambda_2\mathbf{I}$, est respectivement $\{0, \lambda_1 - \lambda_2, \lambda_3 - \lambda_2\}$. Nous pouvons vérifier le signe du produit de ses valeurs propres avec l'expression de l'équation 4.5 :

$$(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2) = -6\sqrt{3}p^2 \sin \alpha (\cos \alpha - \frac{\sqrt{3}}{3} \sin \alpha) \quad (4.9)$$

Comme $\cos \alpha - \sqrt{3}/3 \sin \alpha$ varie dans l'intervalle $[0, 1]$ pour $\alpha \in [0, \frac{\pi}{3}]$ l'expression apparait clairement négative. Ceci valide que λ_2 est bien la valeur médiane des trois valeurs propres de (Q, ψ) . Nous pouvons également constater que $\lambda_2 - \lambda_3 \geq 0$ car $\sin \alpha \geq 0$ pour $\alpha \in [0, \frac{\pi}{3}]$ —ceci confirme donc que $\lambda_1 - \lambda_2 \geq 0$ — ce qui nous laisse :

$$\lambda_1 \geq \lambda_2 \geq \lambda_3$$

En utilisant le rappel 4.2.1 et le fait que la signature est un invariant projectif, nous pouvons confirmer que la signature de $L_2 = C - \lambda_2 \omega$ est bien la même que celle de $E - \lambda_2 I$ c'est-à-dire (1,1) et qu'il s'agit de l'unique conique possédant cette signature. \square

On rappelle que L_2 est la conique contenant \mathbf{v}_∞ et \mathbf{v}_o , qui sont deux droites a priori indistinctes qu'on note $\mathbf{v}_{1,2}$.

Puisque la matrice L_2 est réelle, symétrique, d'ordre 3 et de rang 2, on peut obtenir sa décomposition en vecteurs propres de la façon suivante :

Proposition 4.2.5. *La décomposition en vecteurs propres de la matrice de la conique de rang 2 peut s'écrire à l'aide des vecteurs des images des points de base $\mathbf{z}_1, \mathbf{z}_3 \in \mathcal{R}^3$:*

$$D_2 = \begin{bmatrix} \frac{\mathbf{z}_1}{\|\mathbf{z}_1\|} & \frac{\mathbf{z}_3}{\|\mathbf{z}_3\|} \end{bmatrix} \begin{bmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & \lambda_3 - \lambda_2 \end{bmatrix} \begin{bmatrix} \frac{\mathbf{z}_1^\top}{\|\mathbf{z}_1\|} \\ \frac{\mathbf{z}_3^\top}{\|\mathbf{z}_3\|} \end{bmatrix} \quad (4.10)$$

On peut finalement avoir l'expression des vecteurs des deux droites :

$$\mathbf{v}_{1,2} = \sqrt{\lambda_1 - \lambda_2} \frac{\mathbf{z}_1}{\|\mathbf{z}_1\|} \pm \sqrt{\lambda_2 - \lambda_3} \frac{\mathbf{z}_3}{\|\mathbf{z}_3\|} \quad (4.11)$$

Démonstration. À noter que ce résultat n'est vrai que pour $K = I$ (ou de façon équivalente $\omega = I$). Dans le cas calibré, on peut toujours s'y ramener en appliquant K^{-1} aux vecteurs des points du plan manipulés. Dans ces conditions $E = C$ et on peut exprimer la décomposition en vecteurs propres de $E - \lambda_2 I$ avec les vecteurs de la base orthogonale, $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$. La décomposition en paire de droites s'obtient de la décomposition par bloc de la matrice en se basant sur l'orthogonalité de \mathbf{z}_1 et \mathbf{z}_3 . \square

Les deux solutions pour le vecteur de la normale du plan de support sont finalement $\mathbf{N}_{1,2} = K^\top \mathbf{v}_{1,2}$ dans le repère caméra, et l'équation (4.11) décrit l'ambiguïté de pose du plan, cf. [10].

4.2.2 Une observation intéressante sur les points de base

Maintenant que nous avons formulé les deux solutions pour la ligne de fuite, nous pouvons faire une observation intéressante sur les points de bases. Ce deuxième résultat clé énoncé ci-dessous constitue en fait la fondation de notre approche pour déterminer des configurations où l'ambiguïté peut être levée.

Proposition 4.2.6. *La droite \mathbf{l}_o de D_2 sépare les points de bases \mathbf{x}_1 et \mathbf{x}_3 sur le plan de support alors que la droite \mathbf{l}_∞ ne les sépare pas. Ainsi, si on note $\bar{\mathbf{x}}$, le vecteur normalisé tel que $\bar{\mathbf{x}} = \mathbf{x}/x_3$, les inégalités suivantes sont vérifiées : $(\mathbf{l}_\infty^\top \bar{\mathbf{x}}_1)(\mathbf{l}_\infty^\top \bar{\mathbf{x}}_3) > 0$ et $(\mathbf{l}_o^\top \bar{\mathbf{x}}_1)(\mathbf{l}_o^\top \bar{\mathbf{x}}_3) < 0$.*

Ces deux inégalités restent vraies après toute transformation affine, attention elles ne sont cependant pas vraies après une transformation projective quelconque.

Comment les conditions de la proposition 4.2.6 peuvent-elles être utiles pour étudier l'ambiguïté de la pose du plan projeté sur notre image ? Peut-on formuler un corollaire dans le plan image, qui sous certaines configurations nous garantit que la droite \mathbf{v}_o de $C - \lambda_2\omega$, image de \mathbf{l}_o , sépare toujours les points \mathbf{z}_1 et \mathbf{z}_3 , images des points de bases \mathbf{x}_1 et \mathbf{x}_3 , alors que l'autre ne les sépare pas ?

Autrement dit si on connaît a priori le signe de $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3)$, peut-on garantir que $\text{signe}((\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3)) = -\text{signe}((\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3))$.

Si ceci est vrai, comme les vecteurs de l'image des points de bases $\{\mathbf{z}_i, i \in \{1, 3\}\}$ sont directement calculables à partir des valeurs propres $\lambda_j, j \in \{1, 3\}$ dans l'équation (4.5), nous pouvons lever l'ambiguïté en choisissant la "bonne" ligne de fuite, \mathbf{v}_∞ , parmi les deux droites de $C - \lambda_2\omega$.

Nous affirmons la proposition suivante pour ce corollaire dont la preuve directe découle des propriétés de quasi-affinité [30] pour les points de base.

Proposition 4.2.7. *Lorsque \mathbf{x}_1 et \mathbf{x}_3 sont tous deux situés soit devant soit derrière la caméra, c'est-à-dire sur le même demi-plan séparé par le plan principal⁵ de la caméra, nous avons : $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3) < 0$ et $(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3) > 0$. Dans le cas contraire nous avons $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3) > 0$ et $(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3) < 0$.*

La démonstration de la première proposition est présentée sous la forme d'un code de vérification en calcul symbolique `Matlab` fourni en annexe D.1. La seconde découle de la première en utilisant les propriétés de quasi-affinité [30] pour les points de base. Nous invitons toutefois le lecteur à prendre connaissance du nouveau paramétrage proposé dans la section suivante avant de lire ce code. Ce paramétrage simplifie en effet de façon importante les expressions formelles des vecteurs de base.

4.3 Paramétrage minimal pour étudier l'ambiguïté

4.3.1 Description du paramétrage

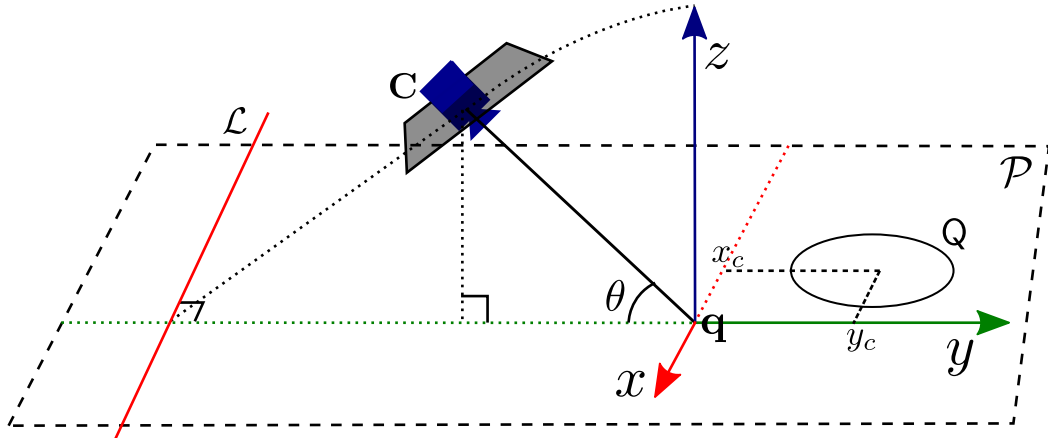


FIGURE 4.4 – Paramétrage proposé pour lever l'ambiguïté.

5. Le plan principal est le plan passant par le centre optique et parallèle au plan image.

Définition 4.3.1. *Nous invitons le lecteur à se référer à la FIGURE 4.4. Soit le point \mathbf{q} , le point du plan \mathcal{P} où passe l'axe optique (ou la droite principale) de la caméra et soit la droite \mathcal{L} , la droite obtenue par l'intersection de \mathcal{P} et du plan principal de la caméra, alors \mathcal{L} est orthogonal à la projection orthogonale de l'axe optique sur \mathcal{P} (on suppose que le plan image n'est pas fronto-parallèle à \mathcal{P}).*

Nous choisissons une représentation euclidienne du monde 3D projectif dans laquelle l'origine est \mathbf{q} tandis que l'axe X est parallèle à la droite \mathcal{L} et l'axe Y est la projection orthogonale de l'axe optique sur \mathcal{P} . Par conséquent, l'axe Z est dirigé par la normale du plan \mathcal{P} , ainsi \mathcal{P} se caractérise par l'équation $z = 0$ dans le repère monde. Nous choisissons l'échelle de telle sorte que le centre de la caméra, de vecteur de coordonnées cartésiennes \mathbf{C} , soit à distance de 1 de l'origine. La FIGURE 4.4 illustre ce nouveau paramétrage.

Soit $\mathbf{C} = [0, -\cos \theta, \sin \theta]^\top$, $\theta \in [0, \frac{\pi}{2}]$, les coordonnées cartésiennes 3D du centre de la caméra où $\pi - \theta$ est l'angle entre l'axe des Y et l'axe optique dans le plan YZ (on retrouve bien une distance de 1 de l'origine). Ainsi la direction de l'axe optique est donnée par $-\mathbf{C}$. Le vecteur de la droite \mathcal{L} a pour coordonnées $\mathbf{d} = [0, 1, \cos \theta]^\top$.

Soit le cercle sur le plan \mathcal{P} de rayon R et centré sur le point de coordonnées cartésiennes $(x_c, y_c, 0)$, ainsi sa matrice représentative dans le plan \mathcal{P} s'écrit :

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ -x_c & -y_c & x_c^2 + y_c^2 - R^2 \end{bmatrix}$$

Précédemment, on a montré que les paramètres des trois membres dégénérés du faisceau de coniques $\{\mathbf{Q} - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ peuvent être calculés à partir des valeurs propres ordinaires de la matrice $\mathbf{T}^{-\top} \mathbf{Q} \mathbf{T}^{-1}$, elles seront notées $\{\lambda_j\}_{j=1..3}$ dans la suite. Si \mathbf{w}_j représente les vecteurs propres de $\mathbf{T}^{-\top} \mathbf{Q} \mathbf{T}^{-1}$ associés aux valeurs propres λ_j , alors les points de bases ont pour vecteurs $\mathbf{x}_j = \mathbf{T}^{-1} \mathbf{w}_j$.

Soit la matrice $\mathbf{Q}' = \sin^2 \theta \mathbf{T}^{-\top} \mathbf{Q} \mathbf{T}^{-1}$, l'ensemble (non ordonné) de ses valeurs propres est

$$\begin{bmatrix} \sin^2 \theta \\ y_c \cos \theta - \frac{1}{2} R^2 - \frac{1}{2} (\tau^{1/2} + x_c^2 + y_c^2 + 1) \\ y_c \cos \theta - \frac{1}{2} R^2 + \frac{1}{2} (\tau^{1/2} + x_c^2 + y_c^2 + 1) \end{bmatrix} \quad (4.12)$$

où τ est une expression positive dépendant de R, x_c, y_c et θ . On peut appuyer la positivité de τ en vérifiant que si la variable τ n'est pas positive la matrice symétrique réelle \mathbf{Q}' ne serait pas diagonalisable.

Maintenant qu'on dispose de l'expression symbolique des valeurs propres, il est possible de vérifier le résultat de la proposition 4.2.6 de la section précédente. Les calculs ont été réalisés avec le moteur de calcul symbolique de `Matlab` dont nous fournissons le code en annexe D.1 pour permettre de vérifier la proposition.

4.3.2 Une première condition nécessaire et suffisante

À partir du nouveau paramétrage proposé et de l'expression symbolique, il est possible d'exprimer une condition sur la projection quasi-affine des points de bases \mathbf{x}_1 et \mathbf{x}_3 . Nous affirmons ainsi la proposition suivante :

Proposition 4.3.1. *Les points de base \mathbf{x}_1 et \mathbf{x}_3 sont situés, sur le plan \mathcal{P} du même côté de \mathcal{L} si et seulement si*

$$\cos \theta (y_c^2 - R^2) (y_c + \cos \theta) + y_c \cos \theta (1 + x_c^2) + x_c^2 + y_c^2 \leq 0 \quad (4.13)$$

Démonstration. Ici aussi, le code de vérification symbolique est fourni en annexe D.2. On note que nous introduisons pour le besoin du calcul un paramètre supplémentaire qui est l'angle de rotation autour du cercle, mais celui-ci n'influence en rien l'expression de la condition nécessaire et suffisante trouvée. □

Ce nouveau paramétrage est à la fois minimal et intuitif. L'expression obtenue est suffisamment simple et compacte pour pouvoir être utilisée dans des applications concrètes où par exemple le suivi de la pose du cercle permettrait de vérifier la condition.

4.3.3 Une condition suffisante exploitable

Dans cette section, l'étude d'une configuration concrète qui exploite la condition trouvée précédemment est proposée.

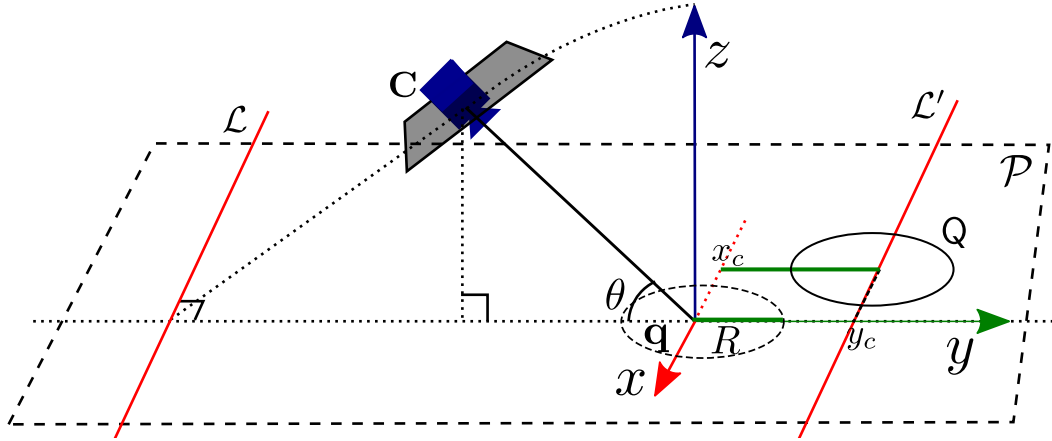


FIGURE 4.5 – Représentation de la condition suffisante pour lever l'ambiguïté.

Soit \mathcal{L}' la droite du plan \mathcal{P} parallèle à \mathcal{L} passant par le centre du cercle comme illustré sur la FIGURE 4.5.

Puisque $\cos \theta > 0$, si nous avons :

$$y_c > 0 \quad (4.14)$$

$$y_c^2 - R^2 > 0 \quad (4.15)$$

alors \mathbf{x}_1 et \mathbf{x}_3 sont situés du même côté de \mathcal{L} . La première inégalité (4.14) stipule que \mathbf{q} doit être du même côté de \mathcal{L}' que la droite \mathcal{L} . La deuxième inégalité (4.15) décrit que le point du plan \mathcal{P} de coordonnées $(0, y_c, 0)$ doit être à l'extérieur du cercle centré sur $\mathbf{q}(0, 0, 0)$ et de même rayon R que Q . Nous sommes alors dans la partie "autrement" de la proposition 4.2.7, la ligne de fuite est donc la droite qui ne sépare pas l'image des points de base. Puisque le point de coordonnées $(0, y_c, 0)$ représente l'intersection de la droite \mathcal{L}' avec l'axe des Y , on peut affirmer qu'une condition suffisante pour résoudre l'ambiguïté est donnée par les deux conditions suivantes :

- (i) \mathbf{q} et la droite \mathcal{L} sont du même côté de \mathcal{L}' ;
- (ii) le point d'intersection de la projection orthogonale de l'axe optique sur \mathcal{P} et de la droite \mathcal{L}' est situé à l'extérieur du cercle centré en \mathbf{q} et de même rayon que Q .

Un exemple de cette configuration peut être donné de la façon suivante : Si on suppose un utilisateur observant avec une caméra un cercle imprimé sur le sol devant



FIGURE 4.6 – Illustration d'une configuration sans ambiguïté

lui, l'ambiguïté est résolue tant que l'utilisateur incline sa caméra vers un point situé entre le cercle et lui et distant de plus d'un rayon du cercle. La FIGURE 4.6 illustre cette configuration.

À noter que cette configuration correspond également à la conditions ou aux conditions d'évaluation des tests que nous proposons en annexe C.

4.4 Conclusion sur l'étude de l'ambiguïté

Dans ce chapitre, nous avons vu comment calculer la pose d'un cercle à partir de son image, en particulier en étudiant l'ambiguïté de la solution à ce problème. Plus précisément, nous avons présenté l'ambiguïté de la solution et son origine dans la résolution algébrique. Nous avons ensuite montré qu'avec un nouveau système de paramètres minimaux (4 paramètres), il est possible de déterminer les configurations ambiguës à l'aide de l'étude d'une expression simple. Dans une configuration illustrée et simple à mettre en place pratiquement, nous avons illustré le fait qu'il est possible de résoudre l'ambiguïté en vérifiant seulement deux contraintes géométriques. Nous pensons que ce travail peut être utile dans des applications spécifiques où la caméra peut être suivie, comme en robotique, par exemple. La contrainte proposée pourrait alors être vérifiée automatiquement pour résoudre l'ambiguïté. Dans ce travail, nous nous sommes intéressés uniquement à la pose du cercle, or, la rotation autour de son axe dans le cadre d'un marqueur est toujours inconnue. Il reste donc à obtenir cette rotation pour obtenir la pose de notre caméra par rapport à un marqueur circulaire. Le chapitre suivant va nous permettre d'aborder cette question.

Chapitre 5

Localisation d'un marqueur enrichi de son environnement

Sommaire

5.1	Combiner marqueurs et points naturels pour faciliter le calcul de pose et la reconnaissance	80
5.1.1	Motivation	80
5.1.2	État de l'art	82
5.2	Estimation de la pose du marqueur avec les correspondances de points	84
5.2.1	Formulation du problème	84
5.2.2	Résolution du problème aux moindres carrés	88
5.3	Distribution des angles estimés à partir des correspondances	91
5.3.1	Résolution directe à partir d'une correspondance	91
5.3.2	Identification de la référence à partir de la répartition des angles	93
5.4	Résolution du problème avec RANSAC (RANdom SAMple Consensus)	95
5.4.1	Méthode d'estimation robuste	95
5.4.2	Évaluation par distance à la droite épipolaire	96
5.5	Résultats	97
5.5.1	Description des tests	97
5.5.2	Résultats sur la reconnaissance des marqueurs	98
5.5.3	Résultats sur l'estimation de la pose	99
5.6	Conclusion et intérêt de la méthode proposée	100

5.1 Combiner marqueurs et points naturels pour faciliter le calcul de pose et la reconnaissance

5.1.1 Motivation

Dans les chapitres précédents, nous avons étudié le calcul de la pose du cercle à partir de son image. Nous avons vu un cadre théorique pour l'étude de l'ambiguïté ainsi qu'une méthode pour travailler avec une caméra calibrée par défaut. Lorsque les marqueurs sont utilisés dans des scènes plus larges, comme nous avons vu précédemment, nous arrivons rapidement aux limites des capacités de reconnaissance de ces marqueurs. Il est donc nécessaire de mettre en place de nouveaux outils pour estimer la pose dans ces conditions. Ainsi, ce nouveau chapitre concerne l'étude du problème du calcul de la pose de la caméra, dans une scène, à l'aide de plusieurs marqueurs en s'aidant de points d'intérêt dans l'environnement. Deux nouveaux problèmes apparaissent alors :

1. Il faut passer du calcul de la pose d'un cercle au calcul de la pose d'une caméra.
2. Il est nécessaire de pouvoir identifier un marqueur à partir de son environnement.

L'estimation de pose de caméra dans une scène large est un problème important à résoudre dans de nombreuses applications, qu'il s'agisse de localisation de robot ou de reconstruction de scènes en 3D. Nous nous plaçons dans un cadre où la scène en 3D n'est pas connue. Les méthodes ont alors pour objectif de retrouver simultanément la structure de la scène et de s'y repérer. Ce problème très étudié porte le nom de *SfM* pour l'anglais *Structure From Motion* dans la communauté de vision par ordinateur et de *SLAM* [57] pour *Simultaneous Location And Mapping* dans la communauté de robotique. Les deux appellations résolvent le même problème et diffèrent, essentiellement, par les communautés concernées et, quelque peu, par la chaîne de traitement mise en place pour sa résolution. L'idée générale est d'utiliser une série d'images prises à différents emplacements de la scène et d'y trouver des correspondances afin de retrouver simultanément la structure de la scène et la pose des prises de vues. Trouver des correspondances entre les images consiste à retrouver, dans plusieurs images, les projections d'une même primitive 3D (point, cercle, droite, etc.). La chaîne de traitement peut se décomposer ainsi :

1. Calculer entre chaque paire d'images des correspondances (nous parlons d'appariement) ;
2. Calculer les poses relatives à partir des correspondances ;
3. Établir un graphe entre les points de vue ;
4. Compléter le graphe en calculant les poses des caméras (*Motion averaging*) ;
5. Optimiser les poses des caméras et la position des points 3D mis en correspondances (nous parlons d'ajustement de faisceaux).

Wu *et al.* [82] proposent une revue et une classification de ces différentes méthodes. Dans certains cas, par exemple dans l'industrie, il est possible de modifier la scène afin d'y ajouter des marqueurs fiduciaux. Les étapes vues précédemment peuvent ainsi être modifiées afin de prendre en compte cet aspect. Munoz *et al.* proposent ainsi d'utiliser uniquement les marqueurs pour reconstruire une carte 3D de la scène [56]. Les auteurs proposent de créer un graphe de pose des marqueurs qui remplace le graphe de pose des caméras dans les méthodes classiques de *SfM*. Celui-ci est établi à partir d'un graphe initial composé des meilleurs arrêtes (pose relative entre marqueurs) calculé sur la base de la plus faible erreur de reprojection sur les images. Enfin, une nouvelle méthode

5.1. COMBINER MARQUEURS ET POINTS NATURELS POUR FACILITER LE CALCUL DE POSE ET LA RECONNAISSANCE

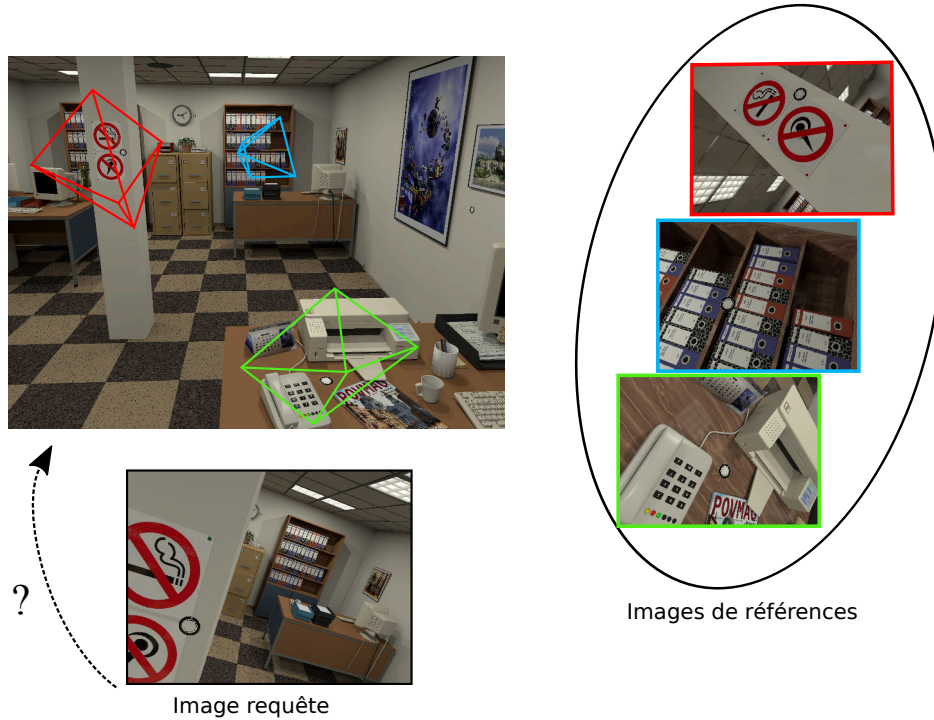


FIGURE 5.1 – Recherche de la pose associée à l'image requête à l'aide d'images de référence – Nous supposons que nous disposons d'images de référence des marqueurs, comme dans le cercle noir à droite. Puis, nous souhaitons reconnaître des marqueurs dans une image requête, comme l'image en bas.

d'ajustement de faisceaux avec l'ajout d'une contrainte sur la position relative des 4 coins d'un marqueur rectangulaire [23] permet de réduire le nombre de paramètres de l'optimisation.

Des travaux portent également sur l'utilisation simultanée des marqueurs et des points naturels, Degol *et al.* proposent d'améliorer le *SfM* en ajoutant des marqueurs fiduciaux dans la scène [41, 16]. L'idée est d'utiliser les correspondances de marqueurs afin d'initialiser de façon fiable le graphe des poses de caméras, de même un ajustement de faisceaux avec la contrainte rectangulaire des marqueurs est utilisé.

Ces travaux montrent l'intérêt d'utiliser conjointement des marqueurs et des points naturels dans une scène large afin de calculer le mouvement d'une caméra. Celui-ci trouve son intérêt quand une scène n'est pas texturée uniformément, il est ainsi possible de compléter certaines zones avec des marqueurs plans. Par la suite, nous allons spécifiquement nous intéresser aux étapes 1) et 2) et notamment à l'utilisation de marqueurs circulaires afin de faciliter le calcul de la pose relative. Ce que nous pouvons constater quand on utilise des marqueurs circulaires dans de grands environnements est qu'il est parfois difficile d'identifier correctement l'intégralité du marqueur. Nous pouvons alors nous demander comment celui-ci peut être identifié dans un jeu d'images. En particulier, nous allons nous intéresser à l'identification et au calcul de la pose entre deux vues d'un marqueur.

Tout d'abord, précisons le cadre simplifié dans lequel nous allons travailler. Si nous disposons d'une scène dans laquelle des marqueurs circulaires ont été disposés, il est possible de prendre préalablement des images de chaque marqueur dans lesquelles celui-

ci apparaît visiblement¹. À partir de ces images, que nous appelons images de référence, une pose précise de la caméra peut être estimée grâce au marqueur. Le problème devient alors celui d'identifier une nouvelle vue correspondant à une nouvelle image, dite image requête. Nous pouvons alors observer deux sous problèmes à résoudre :

1. Identifier le marqueur apparaissant sur l'image ;
2. Retrouver la pose donnant cette vue par rapport au marqueur ou de façon équivalente par rapport à une vue de référence.

Une illustration du problème est donnée sur la FIGURE 5.1.

5.1.2 État de l'art

Extraction et recherche d'image

Dans cette section, nous nous intéressons à résoudre le problème de l'identification du marqueur à partir de la recherche d'image. Cette section ne se veut pas être exhaustive avec les méthodes les plus récentes, mais propose plutôt de donner un contexte global aux méthodes que nous avons utilisées dans la suite des travaux. Le problème de l'identification du marqueur à partir d'une image peut être traité séparément. En effet, nous pouvons rechercher le marqueur de référence, en recherchant dans la base d'images de référence, l'image la plus similaire à notre requête. Cela revient ainsi à résoudre le problème très étudié d'extraction et de recherche d'image dans une base d'images. Les méthodes s'articulent généralement selon la structure suivante [49] :

1. Recherche de points d'intérêts dans l'image, comme l'approche de Harris [29] ou de SIFT, *Scale Invariant Feature Transform* [48] ;
2. Élaboration de descripteurs locaux avec des propriétés d'invariance (Histogramme d'orientations [48], histogramme de gradient [14]) ;
3. Élaboration de descripteurs globaux (arbre de vocabulaire [61]) ;
4. Mise en correspondance (recherche dans un arbre des plus proches voisins) ;
5. Calcul d'un score de similitude.

Nous pouvons noter que récemment, les réseaux de neurones convolutifs ont pris le pas sur beaucoup des méthodes citées précédemment, notamment pour la création de descripteurs locaux [66]. Généralement, ces procédés sont assez longs et fastidieux lorsqu'il est question de mettre en correspondance tous les descripteurs sur un ensemble d'images assez grand. Afin de discriminer rapidement les images de référence qui ne correspondent pas, il est possible de créer des descripteurs globaux à partir de plusieurs descripteurs locaux. Nister *et al.* proposent ainsi de clusteriser dans un arbre les descripteurs locaux à partir de grandes bases de données d'images [61]. Cette structure appelée arbre de vocabulaire permet de caractériser une image en observant la répartition de ses descripteurs locaux dans celle-ci. Ainsi les images peuvent être directement comparées à partir de l'histogramme de la répartition de leurs points dans l'arbre. Cette méthode permet ainsi facilement de discriminer des paires (images requêtes, images de référence) qui auraient trop peu de similitudes.

Globalement, cette première étape présente l'intérêt d'établir une première sélection de candidats pour la recherche de l'image de référence. On peut ensuite utiliser les correspondances des descripteurs locaux trouvées pour calculer la pose.

1. Cela signifie que le marqueur apparaît dans des conditions idéales pour le détecter et le reconnaître.

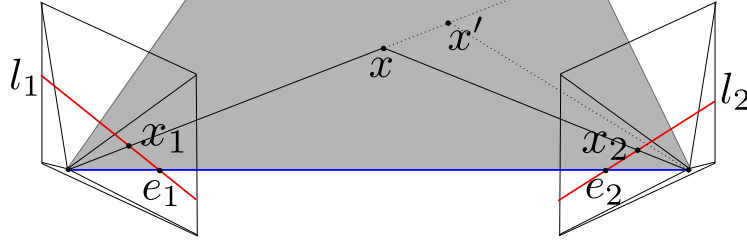


FIGURE 5.2 – Géométrie épipolaire entre deux vues

Calcul de la pose relative sans structure

À partir de correspondances de points dans deux images, il est possible de calculer la pose relative entre les deux points de vues. Cela peut être fait sans structure, c'est-à-dire sans avoir à estimer la position dans l'espace des points 3D correspondants. Ces approches s'appuient sur la géométrie épipolaire qui étudie la géométrie projective intrinsèque entre deux vues, la FIGURE 5.2 l'illustre. La matrice fondamentale notée $F \in \mathbb{R}_{3 \times 3}$ est une matrice dégénérée de rang 2 qui encapsule cette géométrie intrinsèque. En effet, toutes paires de points $(\mathbf{x}_1, \mathbf{x}_2)$ correspondants aux images respectives d'un point 3D observé par deux vues respectent :

$$\mathbf{x}_1^T F \mathbf{x}_2 = 0 \quad (5.1)$$

Cette égalité permet ainsi sans avoir à calculer la position du point 3D, d'obtenir des contraintes sur la géométrie de la scène.

Il est possible de calculer la matrice fondamentale à partir de correspondances de 8 ou 7 points. La méthode à 8 points calcule les coefficients de la matrice F par les moindres carrés ordinaires et recherche ensuite la matrice dégénérée de rang 2 la plus proche (au sens de la norme de Frobenius) de cette première estimation. L'autre méthode consiste à estimer la solution non linéairement mais avec seulement 7 correspondances. Une famille linéaire de solutions de type $F = F_1 + \lambda F_2$ est trouvée et 3 solutions dégénérées correspondant bien à une matrice fondamentale peuvent être extraites.

Dans le cas calibré, si l'on suppose les deux matrices de paramètres intrinsèques connues ou de façon équivalente $K_{1,2} = I$ la matrice fondamentale est alors équivalente à la matrice essentielle E . Cette matrice a la même propriété et peut s'écrire à partir de la pose relative d'une caméra par rapport à une autre, ainsi :

$$E \sim [T]_{\wedge} R \quad (5.2)$$

Retrouver la pose relative peut être alors fait directement à partir de la matrice essentielle. Nister *et al.* [60] proposent d'utiliser ainsi 5 correspondances de points pour le faire. L'idée consiste à extraire un système d'équation polynomiale à 3 variables définies à partir de la matrice essentielle. La résolution donne 10 solutions pour la pose qu'il est nécessaire de distinguer avec plus de correspondances. Une approche utilisant RANSAC est proposée pour estimer la pose relative à partir de n correspondances de façon robuste. On peut noter que lors de la décomposition de la matrice essentielle, le vecteur de translation n'est trouvé qu'à un facteur d'échelle près. Cependant dans le problème que nous cherchons à résoudre, nous disposons de la connaissance de l'échelle si on connaît la taille du cercle observé, ainsi en ce sens, nous disposons d'une certaine forme de structure de la scène.

Calcul de la pose absolue avec structure

Lorsque l'on considère la structure de la scène, il est possible de calculer la pose de la caméra à partir de points 3D dont la position a pu être estimée, c'est le problème connu sous le nom de *Perspective from n Points*, abrégé *PnP*. Le problème minimal, c'est-à-dire avec le minimum de correspondances est connu sous le nom de P3P et consiste à récupérer la pose d'une caméra calibrée à partir de $n = 3$. Celui-ci donne jusqu'à 16 solutions réelles [42].

Dans le contexte particulier de points appartenant à un même plan, il est possible d'utiliser la méthode décrite par Collins *et al.* [13]. Cette méthode consiste à estimer l'homographie localement sur le barycentre des points 3D où l'homographie peut être estimée avec le moins d'incertitude. Cette homographie permet ensuite de calculer les paramètres de la pose du plan en résolvant une équation aux dérivées partielles. L'équation est obtenue en identifiant l'homographie estimée appliquée en un point à sa projection 3D par la caméra à l'ordre 0 et à l'ordre 1. Nous obtenons ainsi 6 équations qui contraignent la pose du plan. Cette méthode différentielle de l'estimation de la pose s'apparente également à celle proposée par Köser *et al.* [44] utilisant les transformations affines locales obtenues par la comparaison des descripteurs de points d'intérêt. Ces transformations affines locales sont ensuite identifiées à la jacobienne de la projection de ce point par la caméra, ce qui donne également des contraintes sur les paramètres de la pose.

Plus récemment, une approche hybride propose d'utiliser à la fois les correspondances 2D-2D et 2D-3D pour combiner les avantages des méthodes avec et sans structure [8]. L'idée consiste à formuler une série d'équations polynomiales à partir des correspondances 2D-2D et 2D-3D et de les résoudre avec des outils automatisés utilisant les bases de Gröbner [45].

Dans le problème que nous cherchons à résoudre, il est important de noter que l'apport du cercle offre une autre contrainte qu'il est possible de prendre en compte. À notre connaissance, la question de l'utilisation de vues multiples de cercles en combinaison avec des correspondances 2D ponctuelles naturelles a été étudiée de façon marginale. Kahl *et al.* [37] proposent d'estimer la matrice fondamentale en utilisant des correspondances de 5 points et une conique, menant à 10 solutions. Alvarez *et al.* ont étudié l'estimation de l'homographie d'un plan de l'espace vers le plan image dans les scénarios sportifs en utilisant seulement un cercle avec quelques informations supplémentaires par exemple la ligne centrale et le point central ou une ligne de contact [2]. Huang *et al.* ont proposé un cadre général de fusion de cercles et de points pour estimer la pose. Partant d'une estimation initiale calculée à partir de contraintes géométriques, ils ont optimisé l'erreur de reprojection en résolvant un problème de moindres carrés non linéaires [33]. Toutefois, dans le cas d'un seul cercle combiné avec plusieurs correspondances de points supplémentaires, l'approche proposée manque de robustesse, celle-ci ne prend pas en compte les données aberrantes, ce qui limite son application.

5.2 Estimation de la pose du marqueur avec les correspondances de points

5.2.1 Formulation du problème

Formulation générale

Nous considérons un modèle de caméra à sténopé dont les paramètres intrinsèques sont connus, ou, de manière équivalente et sans perte de généralités, dont la matrice des

paramètres intrinsèques [76] K satisfait $K = I$. Nous disposons d'un ensemble de vues de référence d'une cible circulaire pour lesquelles les poses de la caméra ont été calculées. Du point de vue applicatif, cet ensemble d'images de référence constitue une bibliothèque de marqueurs. Sur chaque cible circulaire de référence, il est possible d'attacher un repère 3D centré sur le cercle à son plan de support. Ce plan a ainsi pour équation $z = 0$ et le centre du cercle est placé à l'origine du repère avec un rayon normalisé à 1.

Pour chaque vue requête, l'ellipse qui correspond à la projection du cercle est supposée correctement détectée. Sur une image requête, $E \in \mathcal{S}_3(\mathbb{R})$ correspond à la matrice d'ordre 3 représentant l'ellipse. Les deux solutions pour la pose du cercle par rapport à la caméra requête sont alors estimées en utilisant les méthodes étudiées dans les chapitres précédents, soit en utilisant celle de Chen *et al.* [10], soit en utilisant notre méthode pour le calcul de la normale (cf. section 3.2.2). Les étapes d'une solution consistent à :

1. Calculer la normale du plan de support (cf. section 3.2.2) ;
2. Calculer l'homographie à l'aide de l'ellipse et de la normale (cf. section 2.2.3) ;
3. Factoriser l'homographie sous la forme $R_q [\mathbf{e}_1 \quad \mathbf{e}_2 \quad R_q^\top \mathbf{t}_q]$ pour obtenir la pose du cercle [74].

Une pose résultante est constituée de l'orientation du plan de support et de la position de l'origine de ce plan. L'orientation est donnée par une matrice $R_q \in \mathbb{SO}_3$. Puisque l'image d'un seul cercle ne contraint pas l'orientation autour de son axe, l'homographie obtenue à la deuxième étape fixe cette rotation arbitrairement. Ainsi les deux premières colonnes de la matrice R_q peuvent être contraintes arbitrairement, mais doivent cependant constituer une base orthonormale des vecteurs du plan dans le repère caméra. Enfin la troisième colonne est le vecteur de la normale du plan. La position de l'origine par rapport à la caméra est donnée par le vecteur cartésien, $\mathbf{t}_q \in \mathbb{R}^3$, il s'agit évidemment de la position du centre du cercle par rapport à la caméra. On notera par la suite $P = [R \mid \mathbf{t}]$, la matrice de projection de la caméra. Dans notre cas, cette caméra observe un cercle de rayon 1 centré à l'origine du repère monde.

On a donc d'une part la matrice de projection d'une caméra de référence $P_{\text{ref}} = [R_{\text{ref}} \mid \mathbf{t}_{\text{ref}}]$ qui est complètement connue. Et d'autre part la matrice de projection d'une caméra requête qui observe un cercle dont la pose n'est déterminée qu'à une rotation près autour de son axe. Pour l'image requête, on note la matrice de projection de sa caméra P_q . Pour prendre en compte la rotation autour de l'axe du cercle, on introduit l'angle θ de sorte qu'on ait $P_q = [R_q R(\theta) \mid \mathbf{t}_q]$ avec :

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Supposons maintenant qu'un point 3D, de vecteur homogène $\mathbf{X} \sim [x, y, z, 1]^\top$, est observé à la fois sur l'image de référence et sur l'image requête. On note les vecteurs homogènes correspondants aux points images respectifs : $\mathbf{u}_{\text{ref}} \sim [u_{\text{ref}}, v_{\text{ref}}, 1]^\top$ et $\mathbf{u}_q \sim [u_q, v_q, 1]^\top$. Le système d'équations suivant peut être écrit en utilisant les projections par les caméras :

$$\begin{cases} \mathbf{u}_{\text{ref}} \sim [R_{\text{ref}} \mid \mathbf{t}_{\text{ref}}] \mathbf{X} \\ \mathbf{u}_q \sim [R_q R(\theta) \mid \mathbf{t}_q] \mathbf{X} \end{cases} \quad (5.3)$$

Ce système peut être résolu en déterminant les inconnues θ , x , y et z à partir de ce système d'équations. À noter que chaque projection apporte seulement 2 contraintes, en

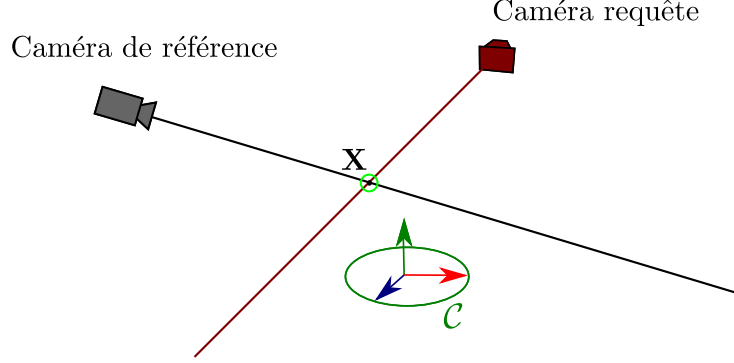


FIGURE 5.3 – Hypothèse d'intersection des rayons provenant des caméras

effet comme nous sommes en coordonnées homogènes, il ne faut pas oublier le facteur d'échelle.

Dans les faits, résoudre le problème pour chaque correspondance de points introduit beaucoup de paramètres inconnus (3 nouveaux pour chaque correspondance). Nous proposons donc de résoudre dans un premier temps le problème sans structure, c'est-à-dire de nous intéresser qu'au calcul de la pose relative de la caméra requête par rapport à la caméra de référence. Autrement dit de calculer le paramètre θ sans calculer la position des points 3D.

Formulation sans structure avec les coordonnées de Plücker

Pour ne pas avoir besoin du point 3D, de vecteur de coordonnées \mathbf{X} , cette hypothèse est utilisée : la rotation $R(\theta)$ de la deuxième caméra (correspondant à l'image de la requête) doit être estimée de telle sorte que les rayons obtenus par la rétro-projection des points images en correspondance \mathbf{u}_{ref} et \mathbf{u}_q se croisent dans l'espace 3D comme présenté sur la FIGURE 5.3.

Pour une caméra dont la matrice de projection est P , la droite 3D correspondant à la rétro-projection d'un point de l'image $\mathbf{u} = [u, v, 1]^\top$ peut être exprimée sous forme de matrice en utilisant la représentation de Plücker :

$$L = P^\top [\mathbf{u}]_\wedge P \text{ et } L = \begin{bmatrix} [\mathbf{m}]_\wedge & \mathbf{d} \\ -\mathbf{d}^\top & 0 \end{bmatrix} \quad (5.4)$$

avec $\mathbf{d}, \mathbf{m} \in \mathbb{R}^3$ où (\mathbf{d}, \mathbf{m}) sont appelées les coordonnées de Plücker de la droite [30, p.72]. On rappelle que les coordonnées de Plücker pour une droite de l'espace projectif représentent les 6 scalaires non nuls d'une matrice antisymétrique d'ordre 4, qu'on note L . Si deux points distincts de vecteurs homogènes \mathbf{a} et \mathbf{b} de l'espace projectif sont sur cette droite, nous avons :

$$L = \mathbf{a}\mathbf{b}^\top - \mathbf{b}\mathbf{a}^\top \quad (5.5)$$

L'intérêt des coordonnées de Plücker réside dans la formulation simplifiée de l'expression de l'incidence entre deux droites de l'espace que nous allons utiliser dans la suite. On

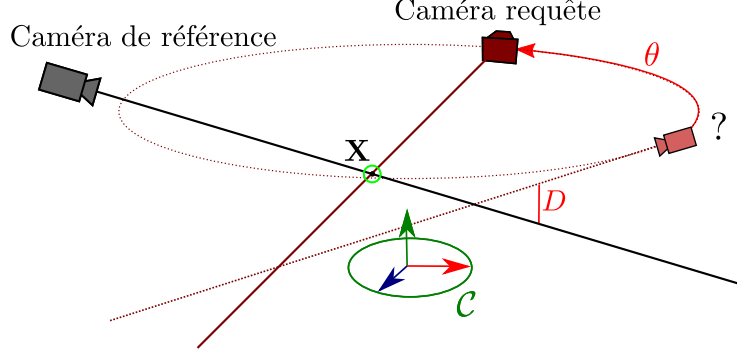


FIGURE 5.4 – La rotation autour de l’axe du marqueur laisse l’image du cercle inchangée. Quel angle θ permet l’intersection du rayon partant de la caméra requête avec celui partant de la caméra de référence ?

note $(\mathbf{d}_{\text{ref}}, \mathbf{m}_{\text{ref}})$, les coordonnées de Plücker de la droite obtenue par rétro-projection d’un point de l’image de référence, \mathbf{u}_{ref} . On peut ainsi écrire :

$$(\mathbf{d}_{\text{ref}}, \mathbf{m}_{\text{ref}}) = ((\mathbf{R}_{\text{ref}})^{\top} \mathbf{u}_{\text{ref}}, \mathbf{c}_1 \wedge \mathbf{d}_1)$$

où $\mathbf{c}_{\text{ref}} = -(\mathbf{R}_{\text{ref}})^{\top} \mathbf{t}_{\text{ref}}$. De façon analogue pour un point \mathbf{u}_q observé sur l’image requête et la matrice de rotation inconnue $\mathbf{R}(\theta)$, on a :

$$(\mathbf{d}_q, \mathbf{m}_q) = (\mathbf{R}(\theta)^{\top} \mathbf{d}'_q, \mathbf{R}(\theta)^{\top} \mathbf{m}'_q) \quad (5.6)$$

$$\text{où } \mathbf{d}'_q = (\mathbf{R}_q)^{\top} \mathbf{u}_q, \mathbf{m}'_q = \mathbf{c}'_q \wedge \mathbf{d}'_q \text{ et } \mathbf{c}'_q = -(\mathbf{R}_q)^{\top} \mathbf{t}_q$$

Pour déterminer la rotation, il nous faut déterminer la valeur de θ pour laquelle la distance entre les deux droites (rayon partant de l’image de référence et rayon partant de l’image requête) soit nulle. C’est-à-dire que les deux droites s’intersectent comme illustré sur la Fig. 5.4. Pour cela, on peut utiliser le *produit réciproque* qui se définit à partir des coordonnées de Plücker [50].

En effet, pour toute paire de droites 3D quelconques, il existe une droite qui intersectent perpendiculairement chacune des droites en deux points. On appelle ces points d’intersection les pieds de la perpendiculaire commune. Un pied est respectivement le point de la droite (qui le contient) le plus proche de l’autre droite de la paire. La distance entre les deux droites est ainsi égale à la distance entre ces deux pieds. Son expression implique ainsi le *produit réciproque* des deux droites qui se définit comme suit :

$$(\mathbf{d}_{\text{ref}}, \mathbf{m}_{\text{ref}}) * (\mathbf{d}_q, \mathbf{m}_q) = \mathbf{d}_{\text{ref}}^{\top} \mathbf{R}(\theta)^{\top} \mathbf{m}'_q + \mathbf{d}'_q{}^{\top} \mathbf{R}(\theta) \mathbf{m}_{\text{ref}} \quad (5.7)$$

L’expression est égal à zéro lorsque les deux droites 3D se croisent. On notera la distance entre les deux droites, D , comme illustré sur la Fig. 5.4. Le problème consiste alors de trouver l’angle θ qui annule cette distance.

Il est également possible d’obtenir une autre interprétation du problème sans structure. En effet si on fait varier θ , la révolution de la droite provenant de la caméra requête

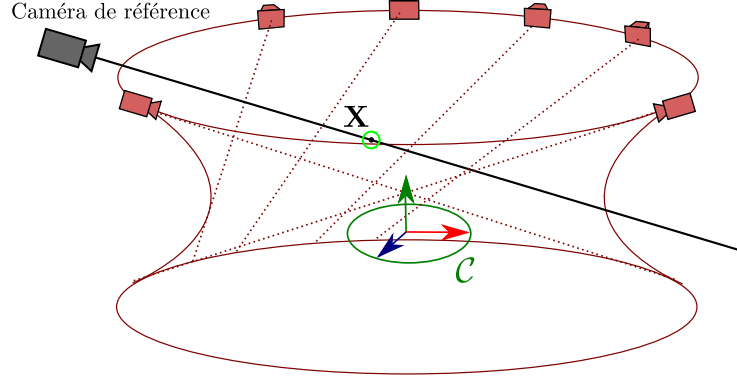


FIGURE 5.5 – Formation de l’hyperboloïde par la révolution du rayon autour de l’axe du cercle

autour de l’axe défini par la normale du cercle \mathcal{C} génère un hyperboloïde à une nappe, voir Fig. 5.5. Le problème revient alors à calculer l’intersection de cet hyperboloïde avec une droite, le rayon partant de la caméra de référence. On peut alors observer qu’il y aura exactement zéro, une ou deux solutions réelles. À noter qu’on écarte la situation où la droite provenant de la caméra requête est parallèle au plan de support. Dans le cas où deux solutions sont possibles, l’ambiguïté ne peut être résolue directement sans autres considérations. À titre d’exemple, dans certains cas théoriques, les intersections de l’hyperboloïde avec le rayon de rétro-projection sont séparées par le plan principal, il serait donc possible d’éliminer une des solutions en supposant que le point est placé devant la caméra. Mais dans les cas pratiques, les deux solutions sont devant la caméra et la plupart du temps dans une petite sphère centrée autour du repère.

La solution correcte pour la pose ne peut être identifiée que si plusieurs points sont impliqués. Dans ce contexte, nous proposons d’étudier des méthodes pour résoudre ce problème avec plusieurs correspondances de points. Nous étudierons d’abord une première solution où le problème est abordé avec une résolution aux moindres carrés. Ensuite nous réaliserons une étude statistique des angles obtenus pour chaque paire de points en correspondance. Cela nous amènera à mettre en place une nouvelle solution robuste aux correspondances erronées en utilisant une approche par consensus aléatoire (RANSAC pour *RANdom SAmple Consensus* [78]). Enfin nous étudierons les résultats obtenus par les deux méthodes proposées.

5.2.2 Résolution du problème aux moindres carrés

Mise en évidence du système linéaire

Comme expliqué précédemment, la révolution du rayon provenant de l’image requête autour de l’axe du cercle forme un hyperboloïde à une nappe. Retrouver la rotation consiste à trouver l’intersection du rayon provenant de la caméra de référence avec cet hyperboloïde. Plusieurs cas sont alors possibles, il y a zéro, une ou deux solutions réelles pour ce problème. Il est facile de voir que l’équation obtenue par l’annulation de (5.7) peut être écrite comme une combinaison linéaire de $x_1 = \cos \theta$ et $x_2 = \sin \theta$ avec la

5.2. ESTIMATION DE LA POSE DU MARQUEUR AVEC LES CORRESPONDANCES DE POINTS

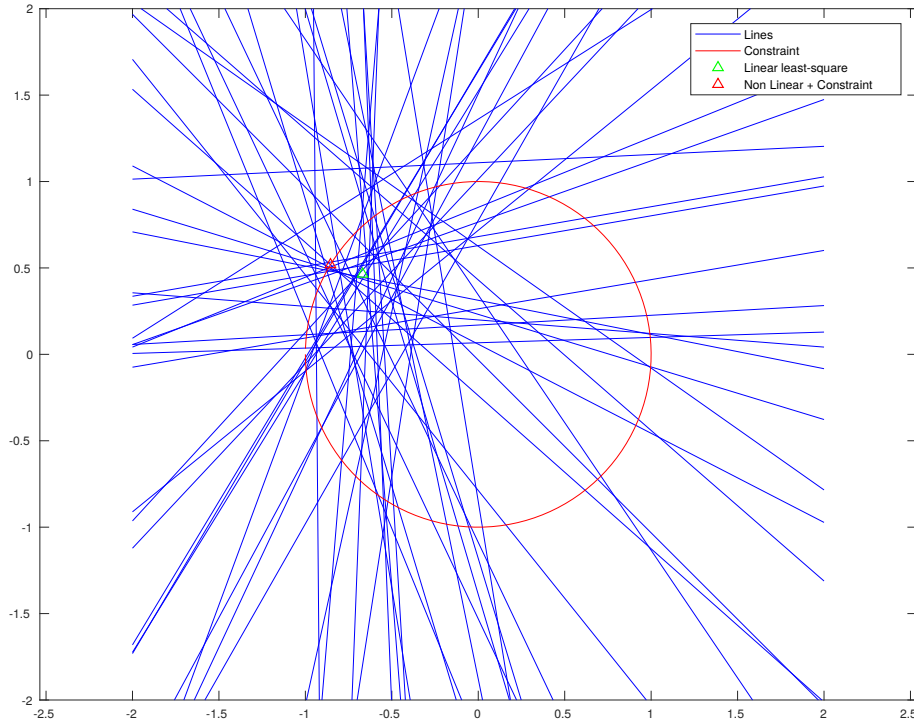


FIGURE 5.6 – Exemple de droites, solution de l'équation 5.9 en x_1 et x_2 .

contrainte $x_1^2 + x_2^2 = 1$. En fait, elle peut être interprétée en 2D comme le problème d'intersection d'un cercle avec la droite paramétrée par (a, b, c) :

$$\begin{cases} x_1^2 + x_2^2 - 1 = 0 \\ ax_1 + bx_2 + c = 0 \end{cases} \quad (5.8)$$

avec

$$\begin{aligned} a &= d_{1,1}m'_{2,1} + d'_{2,1}m_{1,1} + d_{1,2}m'_{2,2} + d'_{2,2}m_{1,2} \\ b &= d_{1,1}m'_{2,2} - d'_{2,1}m_{1,2} - d_{1,2}m'_{2,1} + d'_{2,2}m_{1,1} \\ c &= d_{1,3}m'_{2,3} + d'_{2,3}m_{1,3} \end{aligned}$$

Intéressons-nous maintenant à l'étude de plusieurs paires de points appariés pour la résolution de notre problème. Ainsi au lieu d'avoir l'intersection d'une droite et d'un cercle, nous avons l'intersection d'un ensemble de droites avec un cercle. La Fig. 5.6 donne une illustration dans le plan du nouveau sous-problème. Pour n paires de points en correspondance, on obtient donc un système linéaire du type :

$$\forall i \in \{1 \dots n\}, a_i x_1 + b_i x_2 + c_i = 0 \quad (5.9)$$

auquel il faut ajouter la contrainte circulaire : $x_1^2 + x_2^2 - 1 = 0$.

On peut dans un premier temps résoudre le problème d'optimisation au moindre carré sans contrainte. Si on note \mathbf{A}, \mathbf{C} et \mathbf{X} la matrice et les vecteurs tels que :

$$\begin{aligned} A_{1,i} &= a_i \\ A_{2,i} &= b_i \\ C_i &= c_i \\ \mathbf{X}^\top &= [x_1 \quad x_2] \end{aligned}$$

On a alors le problème d'optimisation au moindre carré linéaire :

$$\min_{\mathbf{x}} \{ \|\mathbf{A}\mathbf{x} - \mathbf{C}\|^2 \} \quad (5.10)$$

En développant et en calculant la dérivée par rapport à \mathbf{x} , on obtient la contrainte d'optimalité :

$$\mathbf{A}^\top \mathbf{A} \mathbf{x} = \mathbf{A}^\top \mathbf{C} \quad (5.11)$$

Si toutes les équations ne sont pas dépendantes, la matrice $\mathbf{A}^\top \mathbf{A}$ est inversible et une première solution peut être trouvée pour x_1 et x_2 en calculant $\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{C} = \mathbf{A}^+ \mathbf{C}$ où la matrice \mathbf{A}^+ est en fait la pseudo inverse de Moore–Penrose [30, p. 144].

Normalisation de l'équation

Il peut être intéressant de normaliser l'équation (5.7) avant de l'utiliser avec un solveur. En effet, si l'on normalise les coordonnées de Plücker des droites 3D, le *produit réciproque* (5.7) correspond au moment entre les droites 3D. On peut alors écrire en utilisant le produit réciproque [50] :

$$\mathbf{d}_{\text{ref}} \mathbf{R}(\theta)^T \mathbf{m}'_q + \mathbf{d}'_q \mathbf{R}(\theta) \mathbf{m}_{\text{ref}} = D \sin \gamma \quad (5.12)$$

où γ est l'angle entre les deux droites et D est leur distance relative. Ainsi, minimiser l'ensemble des équations revient à minimiser la somme des distances entre les rayons rétro-projetés à partir des points de l'image.

Optimisation non linéaire avec la norme L1

La solution obtenue précédemment permet d'obtenir une solution pour x_1 et x_2 mais pas directement pour l'angle θ . On peut revenir à un angle en projetant la solution sur le cercle unité, il suffit de normaliser et de calculer l'angle avec la fonction *atan2*². Le problème de cette solution est qu'elle ne prend pas en compte la contrainte unitaire sur le vecteur $[x_1 \ x_2]^\top$. On peut maintenant chercher une solution directement sur θ par une nouvelle optimisation afin de garantir la contrainte tout en approchant les droites au plus près. Pour cela, nous avons utilisé l'algorithme d'optimisation Quasi-Newton³ [73] pour minimiser :

$$\operatorname{argmin}_{\theta} \sum_{i=1}^n \|a_i \cos \theta + b_i \sin \theta + c_i\| \quad (5.13)$$

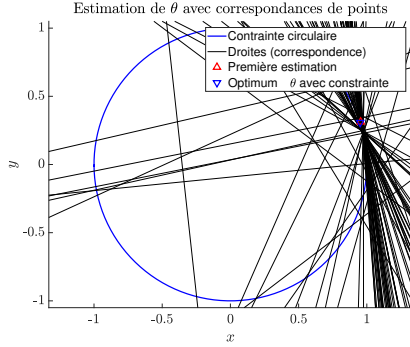
On constate cependant que l'optimisation échoue à trouver un optimum satisfaisant lorsqu'il y a des points mal appariés comme dans les deux exemples présentés sur la Fig. 5.7. Pour le premier exemple, sur la FIGURE 5.7a, on voit en effet que l'intersection des droites et du cercle semble apparaître légèrement plus basse que l'optimum trouvé. Une première solution consiste à utiliser une norme L1 à la place de la norme L2 du moindre carré, cela afin d'être plus robuste aux données aberrantes. Nous minimisons donc maintenant :

$$\operatorname{argmin}_{\theta} \sum_{i=1}^n |a_i \cos \theta + b_i \sin \theta + c_i| \quad (5.14)$$

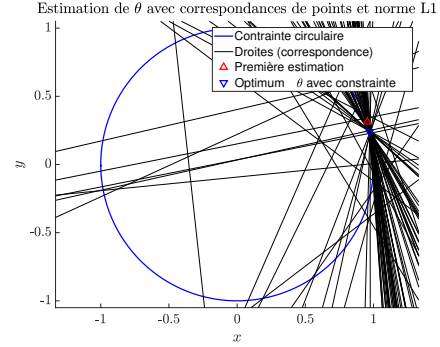
Cela prouve son efficacité sur certains cas, sur la FIGURE 5.7b, l'utilisation de la norme L1 permet de retrouver l'intersection du faisceau de droites.

2. *atan2* est la fonction qui à partir d'un vecteur unitaire 2D donne l'angle de ce vecteur dans le plan
3. L'optimisation a été réalisée avec la fonction *fminunc* de Matlab.

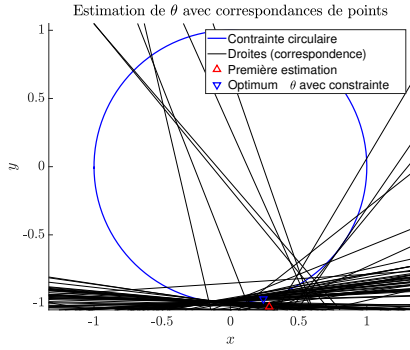
5.3. DISTRIBUTION DES ANGLES ESTIMÉS À PARTIR DES CORRESPONDANCES



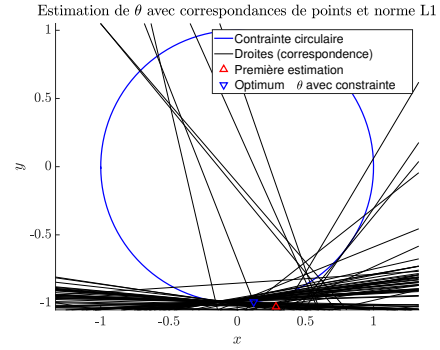
(a) Exemple 1 : Optimisation en norme L2



(b) Exemple 1 : Optimisation en norme L1



(c) Exemple 2 : Optimisation en norme L2



(d) Exemple 2 : Optimisation en norme L1

FIGURE 5.7 – Exemple de droites, solution de l'équation 5.9 en x_1 et x_2

Point sur la méthode proposée

La méthode proposée semble prometteuse pour les cas où il n'y a pas trop de données aberrantes. Cependant dans le second exemple, on peut voir que même avec un estimateur plus robuste l'optimisation ne parvient pas à retrouver l'intersection du faisceau de droites. En effet si la solution robuste est plus proche qu'avec la méthode aux moindres carrés (cf. le triangle bleu inversé sur la FIGURE 5.7c et la FIGURE 5.7d), il semble nécessaire de trouver une autre approche, car celle-ci n'atteint toujours pas l'intersection des faisceaux. L'utilisation d'autres M-estimateurs est possible, mais il semble intéressant d'étudier dans un premier temps la nature de notre problème plus en détail. Dans notre problème, une seule correspondance de points suffit à trouver une ou deux solutions pour les angles, il est donc possible d'étudier plus en détail comment chaque correspondance contribue au calcul de l'angle. C'est ce que nous allons étudier à présent.

5.3 Distribution des angles estimés à partir des correspondances

5.3.1 Résolution directe à partir d'une correspondance

Considérons le problème d'intersection du cercle unité (contrainte $x^2 + y^2 = 1$) et de la droite de coordonnées homogènes $\mathbf{l} = [a \ b \ c]^\top$ correspondant sur le plan cartésien

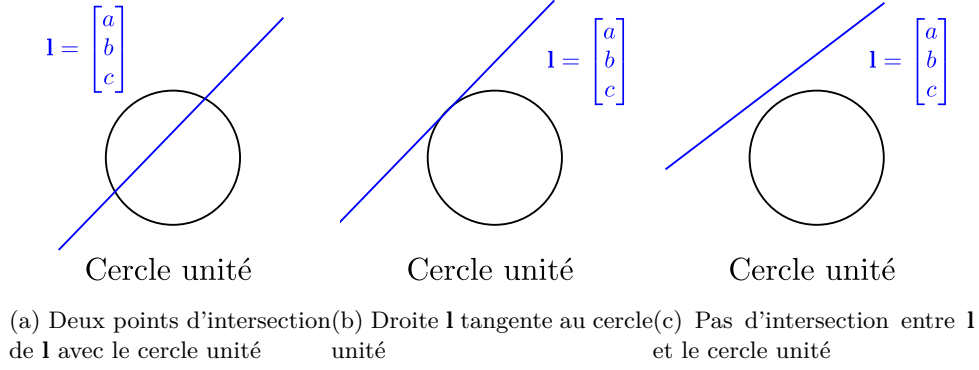


FIGURE 5.8 – Différentes configurations d'incidence entre les rayons

($z = 1$) à la droite d'équation $ax + by + c = 0$. Ce problème modélise bien géométriquement le problème abordé dans la section 5.2.2, lorsque l'on considère uniquement une correspondance. On peut visualiser les trois configurations possibles avec la FIGURE 5.8.

- Première configuration : La droite intersecte le cercle en deux points, il y a donc deux solutions (cf. FIGURE 5.8a).
- Deuxième configuration : La droite est tangente au cercle, il n'y a qu'une seule solution. (cf. FIGURE 5.8b).
- Troisième configuration : La distance entre le cercle et la droite est supérieure au rayon du cercle, il n'y a donc pas de solution. (cf. FIGURE 5.8c).

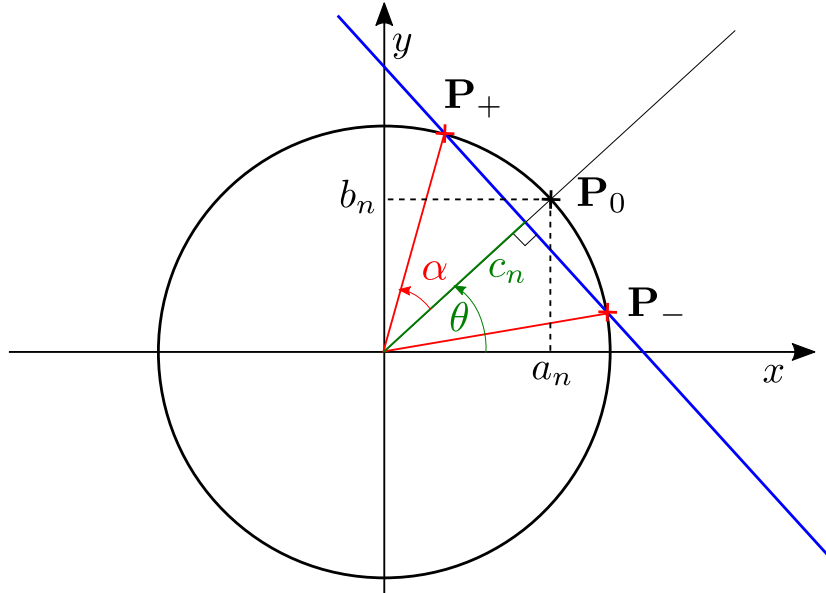


FIGURE 5.9 – Calcul de l'intersection d'une droite et d'un cercle unité

Si nous normalisons l'équation de la droite l par $\frac{\text{signe}(-c)}{\sqrt{a^2+b^2}}$, nous avons pour la droite l'équation :

$$a_n x + b_n y - c_n = 0 \quad (5.15)$$

Remarquons qu'on a alors $c_n > 0$ et c_n représente la distance de la droite l à l'origine du repère. On peut également constater que nous avons : $a_n^2 + b_n^2 = 1$ ainsi il existe θ tel

que $a_n = \cos \theta$ et $b_n = \sin \theta$. θ et c_n sont en fait les coordonnées polaires de la droite.

Nous pouvons constater que le point de coordonnées cartésiennes $\mathbf{P}_0(a_n, b_n)$ est à la fois sur le cercle et sur la droite d'équation $b_n x - a_n y = 0$, une droite orthogonale à \mathbf{l} et passant par l'origine du repère.

Dans le cas où $0 < c_n < 1$, nous avons deux points d'intersection \mathbf{P}_+ et \mathbf{P}_- (cf. FIGURE 5.9). On peut calculer l'angle non orienté α représenté sur la figure en utilisant la fonction *arccos*, ainsi :

$$\alpha = \arccos(c_n) \text{ avec } c_n = \frac{|c|}{\sqrt{a^2 + b^2}} \quad (5.16)$$

Ainsi les coordonnées des points \mathbf{P}_+ et \mathbf{P}_- du cercle unité sont respectivement $(\cos(\theta + \alpha), \sin(\theta + \alpha))$ et $(\cos(\theta - \alpha), \sin(\theta - \alpha))$. Les deux solutions pour l'angle de rotation du problème initial sont donc $\theta - \alpha$ et $\theta + \alpha$ avec :

$$\theta = \text{atan2}(\text{signe}(-c) \frac{b}{\sqrt{a^2 + b^2}}, \text{signe}(-c) \frac{a}{\sqrt{a^2 + b^2}}) \text{ et } \alpha = \arccos(\frac{|c|}{\sqrt{a^2 + b^2}}) \quad (5.17)$$

Dans le cas où $c_n = 1$, nous n'avons qu'un point d'intersection \mathbf{P}_0 , ce point a pour coordonnées $(\cos \theta, \sin \theta)$. La solution pour l'angle de rotation du problème initial est donc $\theta = \text{atan2}(\text{signe}(-c) \frac{b}{\sqrt{a^2 + b^2}}, \text{signe}(-c) \frac{a}{\sqrt{a^2 + b^2}})$.

Dans le cas où $c_n > 1$, il n'y a pas de points d'intersection. Le problème n'admet donc pas de solution par une résolution exacte de l'équation. On peut noter qu'il existe quand même un point du cercle qui est le plus proche de la droite et ce point est le point \mathbf{P}_0 de coordonnées $(\cos \theta, \sin \theta)$.

Cependant le problème de cette résolution directe est qu'elle ne prend pas en compte le bruit ou l'erreur sur le calcul flottant. Ainsi la résolution algébrique amène dans certains cas à deux solutions alors que la distance entre la droite et le cercle est très faible. Dans le cas où il y a deux solutions ceci se manifeste par un écart entre les deux angles solutions, 2α , très faible. La mesure de cette valeur nous permet donc de donner un critère pour préférer une seule solution, l'angle θ , à ces deux solutions, $\theta - \alpha$ et $\theta + \alpha$. On peut généraliser α dans le cas où il n'y a pas d'intersections en calculant $\alpha = \arccos(\frac{1 - |1 - |c||}{\sqrt{a^2 + b^2}})$. Ainsi si $\alpha < \alpha_{\text{seuil}}$ ⁴, nous prenons pour solution θ . Cette méthode permet d'éviter de passer à côté d'une bonne solution approchée s'il n'y a pas d'intersection et de sélectionner une seule solution approchée si les deux solutions algébriques sont trop proches.

5.3.2 Identification de la référence à partir de la répartition des angles

Dans cette section, nous souhaitons observer comment chaque point contribue au choix d'un angle θ et s'il est possible d'étudier la répartition de ses contributions⁵ pour déterminer l'angle. Sur la FIGURE 5.10, nous présentons un exemple de faisceaux et leurs points d'intersection avec le cercle. Nous rappelons que chaque point d'intersection entre le cercle et la droite correspond à une solution pour l'intersection des faisceaux provenant de la correspondance de points entre l'image de référence et l'image requête. Une grande majorité des droites s'intersectent sur un même point du cercle, correspondant à l'angle de la rotation entre les caméras. Cependant, les droites ont tendance à être regroupées autour d'une même forme 3D qui est proche de la forme d'un cône. Cela peut s'expliquer,

4. On pourra prendre une valeur $\alpha_{\text{seuil}} = 1 \text{ deg}$

5. Nous appelons ces contributions, des votes, dans le reste de ce chapitre.

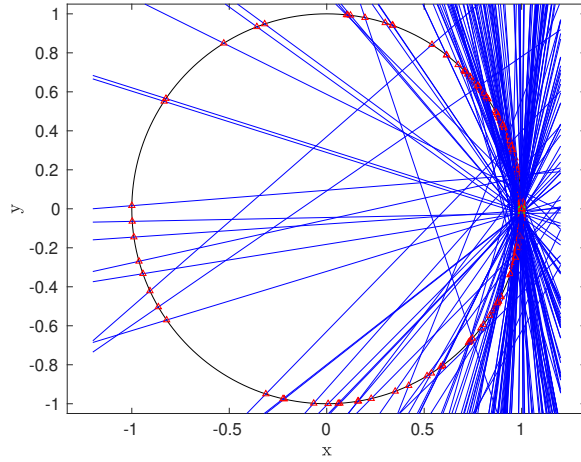


FIGURE 5.10 – Intersection des droites avec le cercle

sur une image réelle, car beaucoup de correspondances sont localisées sur une petite zone texturée de l'image. Les faisceaux sont donc en général regroupés. Le problème de cette distribution est que les autres points d'intersection de ces droites avec le cercle sont donc également proches les uns des autres. Lorsque nous comptabilisons tous les angles où le cercle intersecte les droites dans un histogramme, nous avons la distribution de la FIGURE 5.11. Cette distribution où nous avons classé les angles sur 40 *bins*, nous montre que les angles s'étalent autour de la classe majoritaire mais de façon non centrée. Si nous estimons la distribution avec une loi normale centrée circulaire par exemple la loi de Von Mises (en rouge sur la figure), nous pouvons constater que la distribution obtenue ne convient pas. En effet, l'espérance est légèrement décalée et la variance ne correspond pas à la distribution observée. Ceci remet en question le calcul direct à partir de toutes les correspondances. Cependant, l'existence d'une classe fortement majoritaire permet d'envisager des méthodes robustes qui cherchent à calculer la solution à partir d'un support de données correctes en écartant les données aberrantes.

Si l'image requête et l'image de référence correspondent bien à deux images du même marqueur, et si les correspondances de points obtenues correspondent à des images de points réelles, alors chaque correspondance doit permettre de calculer le même angle de rotation. Si nous considérons ainsi la distribution des angles calculés à partir des correspondances entre chaque image de référence et l'image requête, nous devons observer une distribution bien plus centrée lorsque l'image de référence et l'image requête correspondent à l'image du même marqueur. La FIGURE 5.12 illustre sur un exemple tous les histogrammes obtenus pour chaque image de référence. Nous observons que seul l'histogramme correspondant à l'image de référence observant le même marqueur semble proposer une distribution centrée sur une valeur d'angle avec suffisamment de correspondances. Il est ainsi possible de sélectionner la bonne image de référence en considérant l'image de référence qui donne un vote majoritaire pour un angle donné.

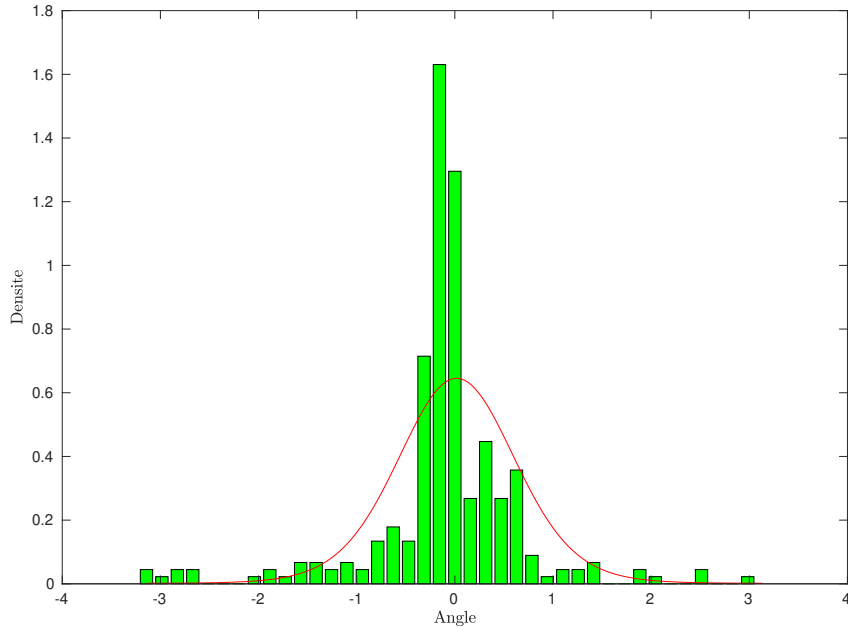


FIGURE 5.11 – Histogramme de distribution des angles solutions

5.4 Résolution du problème avec RANSAC (RANdom Sample Consensus)

5.4.1 Méthode d'estimation robuste

Dans la section 5.2.2, nous avons vu une première proposition qui implique d'utiliser simultanément toutes les correspondances pour calculer une solution. Cependant, dans la pratique, il existe de nombreuses correspondances de points incorrectement appariés, ce qui peut être vu comme des données aberrantes dans le processus de minimisation.

Nous avons vu qu'une seule correspondance suffit pour calculer un angle. En effet, nous pouvons résoudre directement l'équation d'intersection entre la droite et le cercle, ce qui conduit à 0, 1 ou 2 solutions pour l'angle. Nous proposons d'utiliser une approche par consensus fondée sur un échantillon aléatoire, la méthode MSAC (*M-estimateur Sample and Consensus*) [78], une variante de l'algorithme RANSAC. La méthode RANSAC consiste à sélectionner un nombre minimal de correspondances (seulement un dans notre cas) pour estimer l'angle et, ensuite, à évaluer parmi tous les autres points, les points qui respectent le modèle de pose (estimé avec cette correspondance). Dans notre cas, à chaque sélection d'un point, nous avons 0, 1 ou 2 modèles de pose possibles. Le processus est ensuite répété itérativement pour minimiser une fonction d'ajustement tout en s'assurant qu'un nombre suffisant de points valide le modèle. Les points respectant le modèle peuvent ensuite être utilisés à la fin du processus pour estimer l'angle avec plus de précision.

S'il est possible de considérer RANSAC comme un algorithme minimisant une somme sur toutes les données, composée de 0 pour les données correctes et d'une valeur constante pour toutes les données aberrantes, la variante MSAC propose de remplacer cette valeur

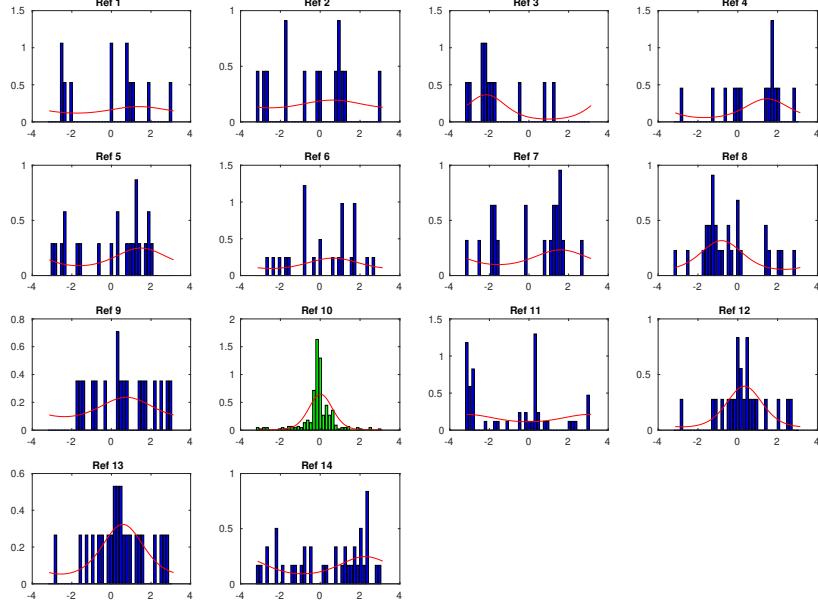


FIGURE 5.12 – Histogramme des angles donnés par chaque correspondance avec leur distribution de Von Mises estimée.

constante dans la fonction coût par une mesure d'erreur robuste sur chaque donnée aberrante. Torr *et al.* [78] proposent ainsi l'utilisation d'un M-estimateur dans la fonction coût de RANSAC, ce qui justifie le nom de la méthode. Les auteurs expliquent que cette méthode peut efficacement se substituer à RANSAC et fournir de meilleurs résultats sans augmenter le coût de calculs pour des données fortement bruitées.

5.4.2 Évaluation par distance à la droite épipolaire

La position réelle du point 3D n'étant pas connue, nous proposons d'évaluer l'erreur en considérant la distance entre un point et la reprojection de la droite épipolaire correspondante dans l'image. La matrice essentielle peut être construite comme suit en utilisant la matrice de projection de la caméra de référence \mathbf{P}_{ref} et la matrice de projection de la caméra de requête \mathbf{P}_q :

$$\mathbf{O}_{\text{ref}} \sim \mathbf{P}_q [\mathbf{t}_{\text{ref}} \ 1]^\top$$

$$\mathbf{E} \sim [\mathbf{O}_{\text{ref}}]_{\wedge} \mathbf{P}_q \mathbf{P}_{\text{ref}}^+$$

où $\mathbf{P}_{\text{ref}}^+$ est la matrice pseudo-inverse de la matrice de projection de la caméra de référence (voir [30, p. 244]). Nous pouvons noter que \mathbf{O}_{ref} est le vecteur représentant le centre de la caméra de référence dans le repère de la caméra requête. Ensuite, pour calculer la distance d'un point au modèle, nous calculons la ligne épipolaire du point correspondant dans l'image requête, $\mathbf{l}_{\text{ref}} \sim \mathbf{E} \mathbf{u}_{\text{ref}}$, puis nous calculons sa distance à cette ligne. Par conséquent, plus la distance est grande, plus il est probable que le point soit une valeur

aberrante. Pour distinguer les valeurs correctes des valeurs aberrantes, nous pouvons appliquer un seuil à cette distance⁶.

5.5 Résultats

Nous avons évalué notre méthode sur des images synthétiques et réelles. Nous calculons la précision de la pose relative de la caméra requête. Et nous évaluons également l'exactitude de l'identification de l'image de référence, c'est-à-dire si notre approche est capable d'identifier l'image de référence correspondante.

5.5.1 Description des tests

Base d'images de tests. Pour les tests synthétiques, nous avons utilisé le bureau virtuel *the office* du dataset ICL-NUIM⁷ pour générer une scène virtuelle en ajoutant des marqueurs circulaires sur différents éléments de la scène, ainsi que des images synthétiques de différents points de vue à l'aide de PovRay⁸. Plus précisément, 8 marqueurs circulaires ont été positionnés dans la scène, à différents endroits. Nous avons d'abord généré une image de référence pour chaque marqueur. Ces acquisitions simulées sont construites avec un point de vue suffisamment proche du marqueur afin de calculer la pose à partir de celui-ci sans nécessiter d'autres points. De même, la lecture de l'identifiant peut se faire sans difficulté. Ensuite, 25 nouvelles images ont été générées à partir de nouveaux points de vue, en général placés dans des configurations où le marqueur est moins identifiable, c'est-à-dire plus éloigné ou placé avec un plus grand angle d'incidence par rapport à la caméra. Ces 25 vues sont utilisées comme images requêtes, c'est-à-dire des images où l'on souhaite reconnaître et localiser le marqueur, à partir de son contour et de points d'intérêt détectés autour.

Pour les images réelles, nous avons placé 14 marqueurs dans un bureau. Notre vérité terrain pour les poses de caméra a été calculée à l'aide d'une mire placée autour de chaque marqueur. Enfin, 44 images du marqueur ont été prises à l'aide d'une caméra de *smartphone* (OnePlus 5) avec au moins 2 vues différentes pour chaque marqueur. Quelques exemples des images synthétiques et réelles utilisées dans les tests peuvent être vus dans la FIGURE 5.13.

Points d'intérêt. Ils ont été détectés et appariés pour chaque paire d'images de référence/requête en utilisant la librairie AliceVision⁹ s'appuyant sur l'opérateur SIFT (*Scale Invariant Features Transform*) [48] pour la détection et l'appariement. De nombreux autres descripteurs ont été proposés dans la littérature, Hassaballah *et al.* en dressent une revue dans [31]. Cependant ce n'est pas l'objet de notre travail et nous avons choisi SIFT car il reste le plus utilisé et le plus connu. Nous avons limité le nombre de points à 1000 par image. Cette restriction garantit que les meilleurs points sont conservés et qu'ils couvrent une grille également répartie sur l'image. Les points ont été appariés et filtrés avec comme critère le rapport de Lowe [48] fixé à 0,8.

Les ellipses dans les images synthétiques et réelles sont détectées en suivant la méthode de Fornaciari *et al.* détaillée dans la section 1.3.2 et en utilisant l'ajustement

6. Après une étude expérimentale, nous avons constaté qu'autour de 5 pixels, les correspondances valides sont bien conservées et les données aberrantes rejetées.

7. <https://www.doc.ic.ac.uk/~ahanda/VaFRIC/iclnuim.html>

8. <http://www.povray.org>

9. <https://alicevision.github.io/>

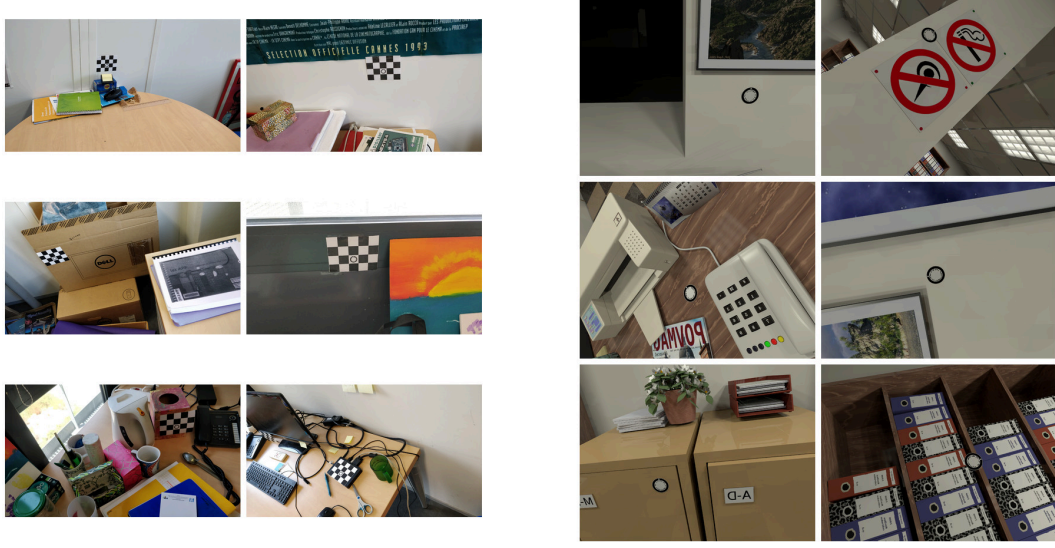


FIGURE 5.13 – Quelques exemples d’images de test : images réelles à gauche, synthétiques à droite.

proposé dans la section 3.3.2.

Critère d’évaluation. Pour chaque image requête, en utilisant la méthode proposée, nous calculons l’identifiant du marqueur et la pose de la caméra par rapport au plan. En ce qui concerne la pose, nous pouvons évaluer l’angle, θ , qui est le seul degré de liberté inconnu, si nous disposons déjà de l’image d’un cercle. Nous pouvons également évaluer l’erreur sur l’orientation et la position de la caméra estimée à partir de l’image du cercle et des correspondances de points. Le premier critère évalue comment l’ajout de nouvelles correspondances de points permet de calculer le degré de liberté manquant, c’est-à-dire, la rotation. Les deux autres critères évaluent la qualité globale de l’estimation de la pose donnée.

5.5.2 Résultats sur la reconnaissance des marqueurs

Le tableau 5.1 affiche les résultats pour le taux de reconnaissance du marqueur et la FIGURE 5.16 montre l’erreur sur la pose estimée. Nous distinguons les deux méthodes $P1CnP$ (Perspective avec 1 cercle et n points) proposées : $P1CnP$ *least square* fait référence à celle décrite dans la section 5.2.2 qui résout le problème directement avec toutes les correspondances et $P1CnP$ *MSAC* à celle présentée dans l’approche robuste en section 5.4 et qui résout le problème avec un schéma RANSAC.

TABLE 5.1 – Taux de reconnaissance des marqueurs : nous distinguons données réelles et synthétiques

Méthode	Précision	
	Réel	Synthétique
$P1CnP$ least-square	70.45 %	88 %
$P1CnP$ MSAC	84.09 %	84 %

Ces résultats montrent que les marqueurs sont généralement bien identifiés à l'aide des images de référence, comme le montre le tableau 5.1. De plus, nous pouvons observer, sur la FIGURE 5.14, un exemple d'identification réussie entre une image requête et une image de référence. Le consensus sur la rotation permet de valider la correspondance des points entre les deux images et ainsi de retrouver la pose relative de la caméra requête par rapport à la pose de la caméra de référence. Cependant, il y a des cas où l'identification échoue, cela survient lorsque de bonnes correspondances peuvent être trouvées sur une autre image de référence que celle correspondant au marqueur que nous cherchons à identifier. Si les points en correspondance sont localisés dans une zone restreinte de l'image, ceux-ci peuvent tout de même trouver un consensus sur une rotation et valider un mauvais modèle de pose relative. La FIGURE 5.15 illustre ce cas. Ainsi, si nous ne considérons que l'identification par la validation du modèle de rotation, en utilisant la proportion de points valides, une mauvaise référence peut parfois être sélectionnée. Dans les faits, cela se produit lorsque de mêmes points d'intérêt sont partagés entre plusieurs images de référence.

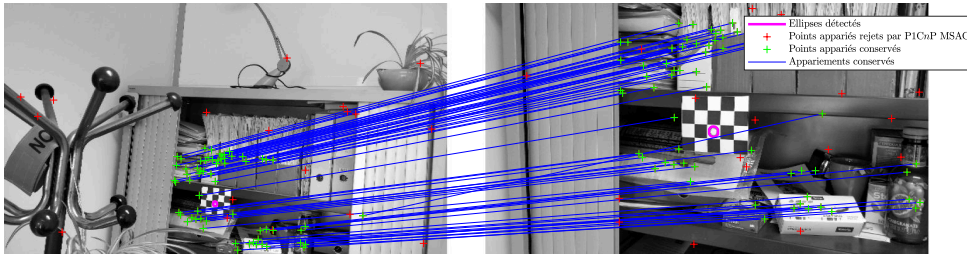


FIGURE 5.14 – Exemple d'images requête (à gauche) correctement identifiées à une image de référence (à droite).

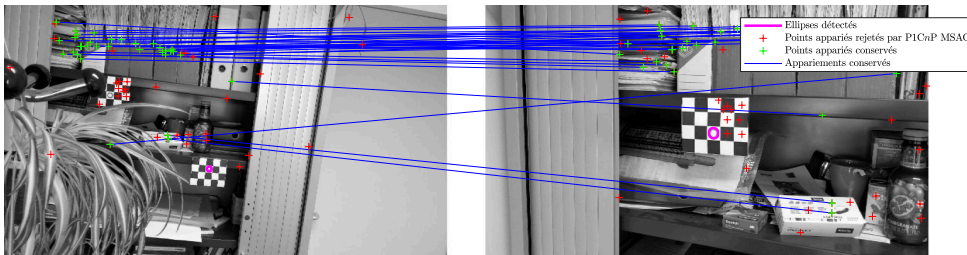


FIGURE 5.15 – Exemple d'images requête (à gauche) incorrectement identifiées à une image de référence (à droite).

5.5.3 Résultats sur l'estimation de la pose

Une première remarque générale est que les résultats sur les premières images synthétiques présentent des erreurs plus grandes que sur les images réelles comme nous pouvons le voir sur la FIGURE 5.16a et la FIGURE 5.16b. Cela peut s'expliquer par le fait qu'il y a plus de régions non texturées ou de textures répétitives dans les images synthétiques par rapport aux images réelles. De ce fait, nous trouvons des régions avec peu de points d'intérêt ou bien des correspondances incorrectes. Par conséquent, l'estimation de la pose contient plus d'erreurs. Ainsi sur la FIGURE 5.16b, nous pouvons voir

que des correspondances de points supplémentaires par rapport aux images synthétiques donnent une estimation précise de la rotation autour de la normale du plan, la médiane est inférieure à 2 degrés pour les images réelles. Sur la FIGURE 5.16a, nous pouvons voir que la méthode utilisant seulement 1 point combinée avec MSAC est plus précise que la méthode directe des moindres carrés avec la contrainte quadratique.

Les autres figures (FIGURE 5.16c, 5.16d, 5.16e, 5.16f) montrent respectivement les résultats de la méthode complète pour estimer la pose : l'orientation et la position. Nous pouvons observer que les résultats sont très satisfaisants pour l'estimation de la pose pour des images réelles surtout si nous considérons la taille relativement petite du marqueur dans l'image, cf. FIGURE 5.13 avec le marqueur au centre de l'échiquier. De plus, nous pensons que la pose obtenue peut être facilement améliorée si nous supposons que nous pouvons utiliser cette estimation initiale dans un processus d'optimisation ultérieure en utilisant les coordonnées des points valides, calculées par triangulation. En ce qui concerne le temps d'exécution de l'estimation de la pose, l'algorithme MSAC converge assez rapidement. Par exemple avec 1000 correspondances entre une image requête et une image de référence, l'algorithme prend environ 30 ms avec *Matlab* et sans optimisation particulière.

5.6 Conclusion et intérêt de la méthode proposée

Dans ce chapitre, nous avons présenté un nouveau contexte permettant d'utiliser une base d'images de référence pour estimer la pose d'un marqueur circulaire à partir d'une image. Nous utilisons des correspondances de points d'intérêt établies grâce à des descripteurs locaux de l'image requête et de l'image de référence pour contraindre la pose du marqueur. Nous supposons également que nous ne connaissons que le contour du marqueur dans l'image requête. L'image d'un contour circulaire ne contraint pas la pose du marqueur entièrement. Il reste en effet à estimer une double ambiguïté et une rotation plane. Ainsi les correspondances de points nous permettent de résoudre l'ambiguïté et d'estimer la rotation avec une approche fondée sur le consensus qui se veut robuste aux mauvais appariements. Cette approche repose sur une validation géométrique qui assure la cohérence des correspondances de points par rapport au modèle de projection. Nous avons évalué notre méthode à la fois sur des données synthétiques et des données réelles. Les résultats obtenus montrent que la méthode est efficace à la fois pour estimer la pose du marqueur dans l'image requête mais aussi pour identifier l'image de référence qui correspond bien au même marqueur parmi les images de la base de référence.

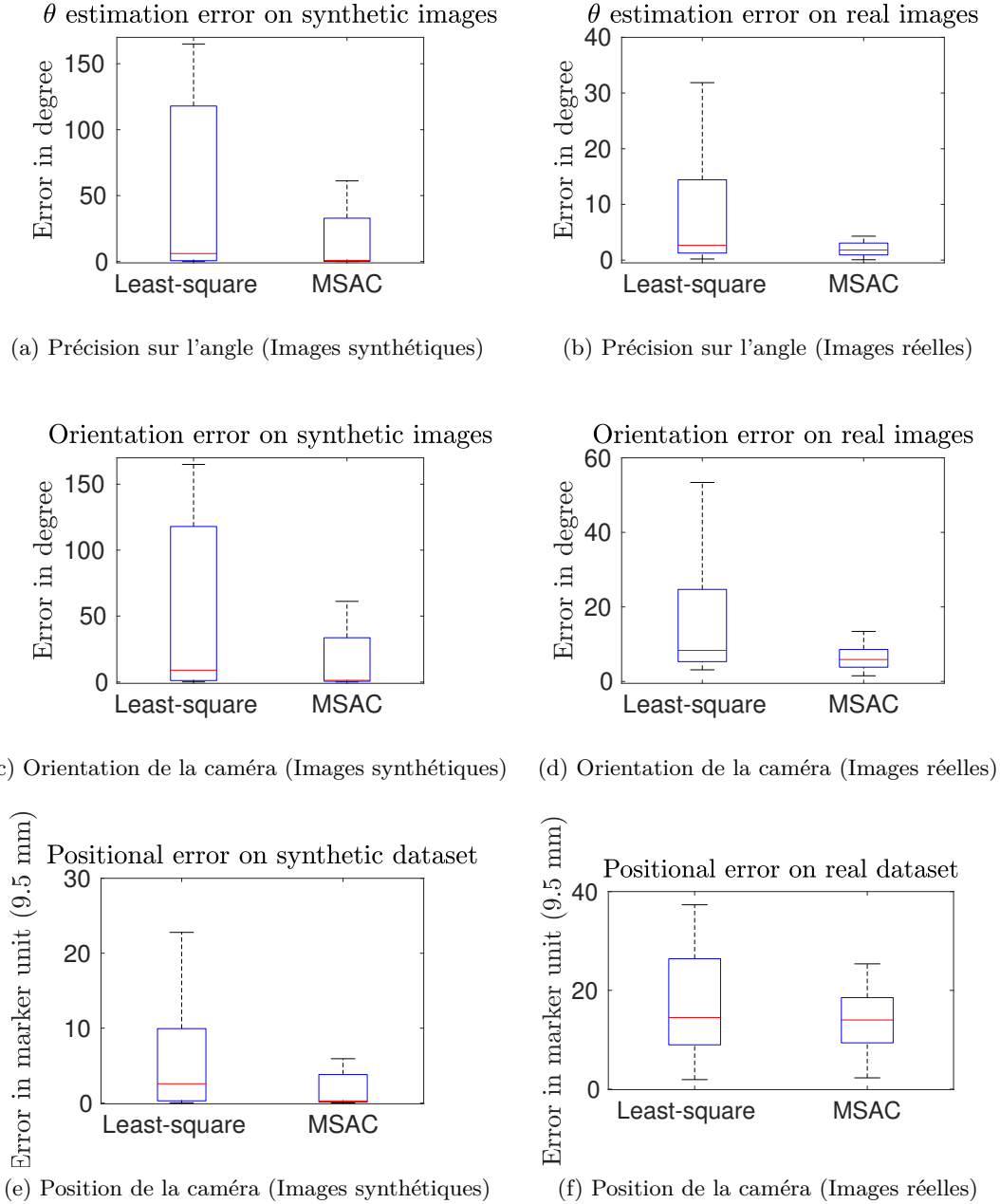


FIGURE 5.16 – Résultats des expérimentations sur images synthétiques et réelles.

Conclusion

Contributions

Dans le chapitre 1, nous présentons un état de l’art sur les marqueurs fiduciaires en distinguant les approches par leurs méthodes pour détecter (dans l’image 2D), localiser dans l’espace et identifier individuellement un marqueur dans sa bibliothèque. Nous constatons qu’il n’existe pas d’approche générique ou de système de marqueurs générique qui peut être utilisé dans tous les cas d’application. De plus, les méthodes dépendent de l’apparence spécifique du marqueur qu’il soit circulaire ou carré ainsi que des critères de performance qui doivent être prioritaires pour l’application visée, comme la fiabilité de l’identification, la rapidité de rejet des faux positifs ou la précision de l’estimation de la pose.

Dans ces travaux, nous étudions les marqueurs spécifiques circulaires, cf. chapitre 2. Les systèmes de marqueurs circulaires existants proposent des approches assez similaires en ce qui concerne l’estimation de la pose. De plus, ce type de marqueur peut être localisé dans l’espace à partir de l’ellipse résultante de la projection sur une image de son cercle contour. À partir de cette ellipse, il est possible de trouver une homographie paramétrant la perspective induite par la projection du plan de support du cercle sur le plan image de notre caméra. Cette homographie n’est cependant pas suffisamment contrainte à partir d’un seul cercle et il est nécessaire d’obtenir d’autres contraintes. Dans l’état de l’art présenté sur cette thématique, nous avons mis en évidence que nous pouvons, soit ajouter d’autres primitives, comme d’autres cercles concentriques, soit estimer les paramètres intrinsèques de la caméra. Ainsi, nous avons étudié et comparé ces deux méthodes avec le marqueur étudié qui est composé de deux cercles concentriques. Ainsi, la première des méthodes utilise l’image du second cercle et la seconde utilise les paramètres intrinsèques de la caméra. Nos résultats montrent que dans les cas où le marqueur est placé à une distance relativement grande de la caméra (taille apparente du marqueur inférieure à 80 pixels), une méthode fondée sur l’utilisation des paramètres intrinsèques est bien plus précise qu’une méthode utilisant les deux cercles. De plus, même en ajoutant l’utilisation d’une valeur approchée de la focale (paramètres intrinsèques), le constat reste le même.

Suite à cette observation, nous développons une expression différentiable du calcul de l’homographie par rapport aux paramètres intrinsèques de la caméra. Cela nous permet dans le chapitre 3 de présenter le concept de calibrage par défaut et de propager les incertitudes sur ces paramètres dans le calcul de l’homographie. Notre méthode propose ainsi de générer des homographies candidates. Nous montrons, expérimentalement, que parmi les homographies proposées, nous pouvons calculer une rectification très proche du cas calibré précisément, au sens de l’erreur de reprojection. Nous avons également proposé des critères photométriques en nous inspirant de critères bien connus de l’état de l’art que nous avons adaptés à notre marqueur, notamment en prenant en compte l’invariance en rotation de l’image du marqueur. Les expériences menées n’attestent cependant pas d’un critère efficace pour sélectionner une meilleure homographie de rectification au sens de l’erreur de reprojection. Ce constat nous amène à la conclusion que l’ambiguïté de rotation doit être calculée antérieurement en utilisant la photométrie de l’image et qu’il paraît difficile de considérer seulement le modèle d’apparence du motif à l’intérieur du marqueur. Ce constat semble d’ailleurs cohérent avec la difficulté à calculer l’homographie avec la méthode utilisant le second cercle interne vue dans le chapitre 2.

En conséquence, dans le chapitre 4 nous proposons d’apporter un nouvel éclairage sur la résolution de l’ambiguïté de la pose du cercle à partir de son image par une projection perspective. Nous reprenons, de manière détaillée, la formulation du chapitre 3 qui nous permet de calculer deux vecteurs solutions pour la ligne de fuite du plan de

support dans l'image. Ces deux solutions algébriques caractérisent l'ambiguïté de pose géométrique de notre problème. Tout d'abord, nous avons observé et démontré qu'il est possible de distinguer ces deux droites solutions. En effet, nous pouvons constater que la véritable ligne de fuite du plan sépare toujours 2 points réels particuliers, les points de base, à condition que ces points soient situés tous deux devant ou derrière la caméra. Au contraire, la seconde droite solution ne les sépare pas. Bien entendu, sans *a priori* supplémentaire, il n'est pas possible de déterminer si les points de base sont du même côté ou non du plan principal de la caméra. Le second apport de ce chapitre est d'étudier, avec un nouveau paramétrage minimal de la pose, une expression formelle qui conditionne la position relative de ces points par rapport au plan principal. Autrement dit si cette expression peut être évaluée, il est possible de résoudre l'ambiguïté de pose du cercle.

Enfin dans le chapitre 5, nous construisons une nouvelle bibliothèque de marqueurs à partir de plusieurs vues de référence de marqueurs placés dans une scène. L'idée est de prolonger le travail amorcé dans le chapitre 3. Cette fois-ci, nous recherchons la rotation du marqueur à partir de correspondances de points entre une nouvelle vue requête et notre bibliothèque de référence. Ces correspondances incluent donc l'usage de points d'intérêt à l'extérieur du marqueur. Afin de rendre notre méthode robuste à l'occurrence de correspondances erronées de points, nous avons développé un algorithme fondé sur une approche de type RANSAC. Notre méthode permet le calcul de la pose du marqueur à partir de la correspondance minimale entre les images d'un point et d'un cercle de la scène entre la vue de référence et la vue requête. Nous montrons également que notre méthode est prometteuse pour identifier le marqueur dans la bibliothèque.

Publications

- **Rectification métrique de l'image d'un marqueur à contour circulaire à partir d'une caméra calibrée par défaut**, Damien MARIYANAYAGAM, Pierre GURDJOS, Sylvie CHAMBON, Florent BRUNET, Vincent CHARVILLAT, ORASIS, 2017
- **Estimation de la pose d'un cercle à partir d'une caméra calibrée par défaut**, Damien MARIYANAYAGAM, Pierre GURDJOS, Sylvie CHAMBON, Florent BRUNET, Vincent CHARVILLAT, Reconnaissance des Formes et Intelligence Artificielle, 2018.
- **Pose estimation of a single circle using default intrinsic calibration**, Damien MARIYANAYAGAM, Pierre GURDJOS, Sylvie CHAMBON, Florent BRUNET, Vincent CHARVILLAT, Asian Conference on Computer Vision, 2018.
- **Camera Localization by Single View Query Using One Circular Target**, Damien MARIYANAYAGAM, Pierre GURDJOS, Sylvie CHAMBON, Vincent CHARVILLAT, Scandinavian Conference on Image Analysis, 2019

Perspectives

Nous avons choisi de présenter les perspectives à court terme dans un premier temps, puis les perspectives à long terme.

Détection du marqueur par de l'apprentissage profond

Nous avons parlé de détection de points d'intérêt et de l'utilisation de ces points d'intérêt, mais nous avons peu discuté sur la façon de choisir un détecteur optimal pour ce genre d'application. De plus, nous n'avons pas présenté et étudié les approches de détection. Ainsi, nous rappelons que les approches sans marqueurs les plus étudiées pour le suivi de la pose d'une caméra suivent les étapes caractéristiques suivantes :

- détection de points d'intérêts ;
- calcul du descripteur des points ;
- appariement des points ;
- calcul du modèle géométrique.

L'émergence de l'apprentissage profond a récemment modifié la conception de ces étapes, à commencer par la construction des descripteurs. Jusqu'alors les descripteurs étaient construits dans l'objectif d'être invariants aux transformations subites par l'objet photographié. Ces transformations se modélisent par exemple sur l'image par la rotation, la translation, le changement d'échelle ou d'intensité. L'apprentissage profond a permis d'automatiser ce procédé en le traitant comme un problème d'apprentissage des caractéristiques invariantes de l'image d'un objet ponctuel. La détection de points est également une étape qui peut directement profiter de l'apprentissage profond. En effet, les points d'intérêt peuvent être appris en considérant les points offrant le plus de répétabilité dans leur détection.

La détection d'un marqueur fiduciaire repose sur la détection de ses primitives particulières. Le problème de détection est ainsi décomposé, le marqueur n'est pas étudié comme un tout, mais comme un ensemble de primitives distinctes. Nous pouvons dès lors nous demander si des méthodes de détection fondées sur l'apprentissage profond ne pourraient pas simplifier la détection des marqueurs dans l'image. Des études récentes dans ce domaine montrent l'intérêt pour ce sujet de recherche [32, 51]. Ce champ de recherche pourrait se révéler pertinent dans le contexte de nos travaux, car il permettrait de s'abstraire des méthodes de détection actuelles qui donnent parfois des résultats peu satisfaisants, en plus d'être coûteuses à implémenter (nous rappelons que dans le cadre de cette thèse nous nous sommes intéressés à l'utilisation des algorithmes sur des appareils mobiles).

Partitionner les poses en fonction de l'ambiguïté liée à la prise de vue

Dans le chapitre 4, nous avons vu qu'il était possible de formuler une expression qui permette de résoudre l'ambiguïté de la pose du cercle en étudiant son signe. Nous pensons qu'en étudiant le signe de cette expression plus en détail sur l'espace des paramètres, nous pourrions partitionner l'espace géométrique en fonction de la solution algébrique à retenir. Cela peut se faire simplement, mais nécessite l'étude rigoureuse de cette fonction à 4 paramètres.

Cette partition pourrait être utile en pratique dans le cadre du suivi de trajectoire. En effet, dans ce cadre, il serait possible à partir des *a priori* sur les poses précédentes d'évaluer le signe de l'expression et son changement probable pour l'acquisition courante. Ces travaux pourraient même être appliqués pour le repérage de robot ou pour la réalité augmentée sur mobile. En effet, la pluralité des capteurs permet d'obtenir de forts *a priori* sur la pose.

Intégrer la méthode proposée dans un algorithme de *SLAM* avec plusieurs marqueurs

Le système de marqueurs proposé dans le chapitre 5 traite les marqueurs indépendamment les uns des autres. Ainsi, la pose d'une caméra est estimée par rapport à un repère local lié au plan du marqueur. Dans nos résultats, nous observons que pour certaines images, des correspondances de points sont trouvées avec l'image de référence d'un autre marqueur. L'apparition de ce problème est due à la présence de plusieurs marqueurs sur une même image requête. Cela signifie que les points d'intérêt sont communs entre les deux images de référence. Ces correspondances erronées ont pour effet de perturber l'identification du marqueur. Pour résoudre ce problème, nous pourrions considérer la construction d'un graphe entre les poses des différents marqueurs de la scène à l'instar de [55] qui propose un système de *SLAM* à partir de marqueurs carrés. Nous pouvons d'ailleurs noter que dans leurs travaux les auteurs doivent considérer le problème d'ambiguïté sur les marqueurs carrés. Ainsi, notre contribution sur cet aspect, pour les marqueurs circulaires, serait complètement pertinente dans cette perspective. Cela nous permettrait d'intégrer notre méthode de calcul de pose local dans un système global de localisation fondé sur des marqueurs circulaires.

Conclusion

Annexes

Annexe A

Rappel de géométrie projective

A.1 Notation

Un vecteur est représenté par un caractère gras, par exemple \mathbf{v} , et une matrice par un caractère sans empattement (de typologie *sans-serif*), par exemple \mathbf{M} . La composante j du vecteur \mathbf{v} sera notée v_j et l'élément (j, k) de la matrice \mathbf{M} sera noté M_{jk} . Le vecteur colonne k de la matrice \mathbf{M} sera noté \mathbf{M}_k et le vecteur ligne j de la matrice \mathbf{M} sera noté \mathbf{M}^j .

Les vecteurs en majuscules contiennent des coordonnées cartésiennes et ceux en minuscules des coordonnées homogènes. Afin de simplifier l'écriture, l'opérateur \sim désignera l'égalité projective. Ainsi, nous notons :

$$\mathbf{v}_1 \sim \mathbf{v}_2 \Leftrightarrow \exists \lambda \in \mathbb{R}^*, \mathbf{v}_1 = \lambda \mathbf{v}_2$$

Nous avons également $[\mathbf{v}]_{\wedge}$ la matrice antisymétrique telle que :

$$\mathbf{v} \wedge \mathbf{w} = [\mathbf{v}]_{\wedge} \mathbf{w}, \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$$

où \wedge est le produit vectoriel de \mathbb{R}^3 .

Enfin, \mathcal{SO}_3 désigne le groupe des rotations dans l'espace tridimensionnel euclidien et $\mathcal{S}_3(\mathbb{R})$ le sous-ensemble des matrices réelles d'ordre 3 formé des matrices symétriques.

A.2 Modèle perspectif de la caméra

Un appareil photographique peut être modélisé par le modèle de la caméra sténopée. Ce modèle autrement appelé caméra à trou d'épingle consiste à considérer un plan image uniquement éclairé par la lumière provenant d'un centre optique (le trou d'épingle). Tous les rayons lumineux venant de l'objet source passent ainsi par ce centre optique et une projection de l'objet apparaît sur le plan image comme sur le schéma de la figure A.1. En réalité, dans le cadre d'un appareil photographique classique, cette convergence des rayons est réalisée par l'utilisation d'une lentille qui a pour effet de focaliser les rayons en un point. Une caméra est donc principalement caractérisée par la distance focale de son système optique, la résolution en pixel de l'image et la position de son plan image (capteur) ainsi que sa forme (rectangulaire en général).

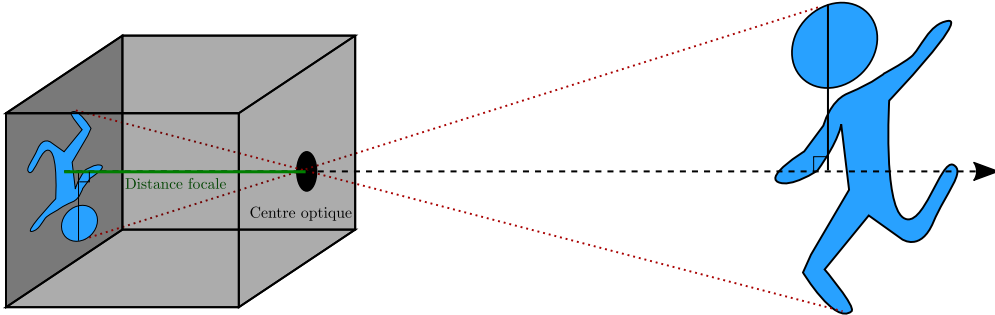


FIGURE A.1 – Modèle de la caméra sténopée

Dans le cadre de nos travaux, les aberrations géométriques et les distorsions optiques générées par la caméra ne seront pas prises en compte. La distorsion la plus communément prise en compte est la distorsion radiale, il s'agit d'une déformation qui translate

les points images dans la direction opposée du point principal et dont l'intensité dépend de l'éloignement du point par rapport au point principal. Il est possible d'approcher cette transformation par une expression polynomiale à l'aide d'un développement limité, les paramètres de l'expression peuvent être estimés au moment du calibrage. Les images peuvent ainsi être rectifiées pour obtenir la projection linéaire que nous décrivons ensuite.

Nous définissons la pose d'une caméra par le couple (\mathbf{R}, \mathbf{T}) dans le repère \mathcal{R} de la scène, avec $\mathbf{R} \in \mathcal{SO}_3$ et $\mathbf{T} \in \mathbb{R}^3$. Les paramètres intrinsèques de la caméra sont représentés par la matrice notée \mathbf{K} .

$$\mathbf{K} = \begin{pmatrix} f_x & \alpha & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{A.1})$$

Ses paramètres sont :

- les distances focales (en pixels) f_x et f_y , selon les axes x et y du plan de l'image ;
- le facteur d'obliquité α ;
- les coordonnées cartésiennes (u_0, v_0) du point principal.

Dans la littérature, nous constatons qu'en pratique le facteur α est souvent très proche de zéro. L'intérêt de cette matrice est qu'elle permet d'exprimer la projection d'un point de façon simple grâce à l'utilisation des coordonnées homogènes. Ainsi un point de l'espace, dont les coordonnées dans le repère \mathcal{R} sont $\begin{bmatrix} x & y & z \end{bmatrix}$, a pour image :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \mathbf{P} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{avec } \mathbf{P} = \mathbf{K} [\mathbf{R} \quad \mathbf{T}] \quad (\text{A.2})$$

A.3 Espace projectif et projection de primitive usuelle

Dans l'espace euclidien, le cinquième postulat des axiomes d'Euclide amène à énoncer que deux droites parallèles non confondues n'ont aucun point d'intersection. Une projection remet en question ce principe. Un exemple de cette remise en question est celui où, au volant d'une voiture sur une route droite, nous percevons que les lignes blanches de sécurité convergent en un point. Cela s'explique par le fait que dans l'image formée sur notre rétine, les lignes blanches s'intersectent réellement. Ainsi, la géométrie projective c'est l'étude d'une géométrie où ce lieu idéal d'intersection existe toujours. Plus formellement, nous pouvons définir un espace projectif comme un espace euclidien augmenté des points d'intersection de toutes les droites de l'espace. Ainsi le plan projectif, \mathbb{P}^2 est composé d'un plan augmenté d'une droite dite à l'infini, notée \mathbf{l}_∞ contenant l'intersection de toutes les droites parallèles. Il est possible d'y définir un nouveau système de coordonnées dites homogènes. Nous représentons alors les points avec une nouvelle coordonnée. Ainsi, pour le plan projectif, nous avons, pour les points qui ne sont pas à l'infini :

$$\begin{bmatrix} kx & ky & k \end{bmatrix}^\top \in \mathbb{P}^2 \quad \text{avec } k \in \mathbb{R}^*$$

Pour les points à l'infini, nous avons :

$$\begin{bmatrix} kx & ky & 0 \end{bmatrix}^\top \in \mathbb{P}^2 \quad \text{avec } k \in \mathbb{R}^*$$

À noter que le point de coordonnées $\mathbf{0}_{\mathbb{R}^3}$ n'est pas un point du plan projectif. Nous pouvons y ajouter la relation d'équivalence \sim , comme introduite dans la partie notation.

A.4 Homographie et projection de plan

Une transformation projective du plan sur lui-même est nommée homographie, qu'on note \mathcal{H} , celle-ci peut être décrite par une matrice 3×3 généralement notée \mathbf{H} qui paramètre une application linéaire sur les vecteurs de coordonnées homogènes. Par souci de simplifier l'écriture, nous confondrons l'homographie et sa représentation en coordonnées homogènes par une matrice \mathbf{H} . Estimer une homographie consiste donc à estimer ses 8 ($3 \times 3 - 1$) paramètres.

Beaucoup de problèmes de vision par ordinateur trouvent leur résolution dans le calcul de l'homographie. Par exemple, la projection d'un plan cartésien vers le plan image d'une caméra par une vue perspective engendre une homographie. Sans perte de généralité, si nous considérons le plan d'équation $z = 0$ et une caméra de pose (\mathbf{R}, \mathbf{T}) , nous pouvons décrire la projection d'un point du plan sur le plan image par la matrice d'homographie $\mathbf{H} = \mathbf{K} [\mathbf{R}_1 \ \mathbf{R}_2 \ \mathbf{T}]$. \mathbf{R}_1 et \mathbf{R}_2 sont respectivement la première et la deuxième colonne de la matrice de rotation \mathbf{R} . Ce résultat est directement obtenu à partir de l'équation de projection d'un point. Ainsi la projection d'un point du plan de coordonnées cartésiennes $(x, y) \in \mathbb{R}^2$ a pour coordonnées homogènes $[u \ v \ w]^\top \sim \mathbf{H} [x \ y \ 1]^\top$. De façon plus générale, il est montré que les images d'un même plan pris par une caméra, mais de deux points de vues différents sont liées par une homographie. De même, une caméra qui serait en rotation autour de son centre optique produit des images toutes liées par une homographie ce qui est généralement utilisé pour construire des mosaïques d'images ou des panoramas.

A.5 Coniques et leurs projections

Une conique peut être définie soit comme l'intersection d'un cône et d'un plan, soit comme les lignes de niveau des fonctions quadratiques. La deuxième définition permet d'explicitier une conique par une équation quadratique sur le plan cartésien du type :

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

L'appartenance d'un point à une conique peut être vérifiée en utilisant la matrice symétrique $\mathbf{C} \in \mathcal{S}_3(\mathbb{R})$ telle que :

$$\mathbf{C} = \begin{bmatrix} A & \frac{B}{2} & \frac{D}{2} \\ \frac{B}{2} & C & \frac{E}{2} \\ \frac{D}{2} & \frac{E}{2} & F \end{bmatrix}$$

Théorème A.5.1. *Ainsi, un point de coordonnées cartésiennes $[x \ y] \in \mathbb{R}^2$ appartient à une conique si et seulement si $[x \ y \ 1] \mathbf{C} [x \ y \ 1]^\top = 0$*

L'intérêt de la forme matricielle est qu'elle permet également d'étudier simplement la nature de la conique en utilisant les propriétés de sa matrice. Toute conique peut être ramenée à une conique du TABLEAU A.1 par une transformation affine.

Définition A.5.1. *La signature projective d'une conique correspond aux couples : $(\max(p, n), \min(p, n))$ avec respectivement p et n , le nombre de valeurs propres positives et négatives de la matrice associée à la conique.*

Ainsi, si nous calculons la signature de la matrice de la conique \mathbf{C} et de sa sous-matrice $\mathbf{C}_{2 \times 2}$, il est possible de connaître sa nature.

Théorème A.5.2. *La signature d'une conique à coefficient dans \mathbb{R} est invariante par transformation projective.*

Démonstration. Cette propriété provient directement de la loi d'inertie de Sylvester. \square

La propriété A.5.2 nous donne la classification projective des coniques (suivant uniquement la première colonne du tableau A.1).

Signature de C	Signature de $C_{2 \times 2}$	Équation	Forme réelle	Forme complexe
(3,0)	(2,0)	$x^2 + y^2 + 1$	vide	ellipse
(2,1)	(2,0)	$x^2 + y^2 - 1$	ellipse	
	(1,1)	$x^2 - y^2 - 1$	hyperbole	
	(1,0)	$x^2 - y$	parabole	
(2,0)	(2,0)	$x^2 + y^2$	1 point	
	(1,0)	$x^2 + 1$	vide	2 droites
(1,1)	(1,1)	$x^2 - y^2$	2 droites sécantes	
	(1,0)	$x^2 - 1$	2 droites parallèles	
(1,0)	(1,0)	x^2	1 droite	

TABLE A.1 – Classification affine des coniques

Théorème A.5.3. *La transformation projective d'une conique C par l'homographie H peut être exprimée comme $H^{-T}CH^{-1}$*

Démonstration. Soit \mathbf{x} , un point sur la conique C ainsi $\mathbf{x}^T C \mathbf{x} = 0$. Soit \mathbf{y} , l'image du point \mathbf{x} par la transformation projective. Nous avons donc $\mathbf{y} \sim H\mathbf{x}$. Nous pouvons constater que $(\mathbf{y}^T H^{-T})C(H^{-1}\mathbf{y}) = 0$ et donc que le point \mathbf{y} réside bien sur la conique $H^{-T}CH^{-1}$. \square

A.6 Quelques propriétés des coniques

Nous pouvons remarquer que dans \mathcal{P}_2 , un point et une droite peuvent être exprimés par les mêmes coordonnées. Cette relation entre ces deux objets s'appelle la dualité. Celle-ci s'applique aussi à d'autres objets du plan projectif et notamment aux coniques. Nous pouvons ainsi définir une conique par l'ensemble de ses droites tangentes.

Définition A.6.1. *Ainsi la conique duale de la conique de points, C, est la conique par laquelle passent les points duals des tangentes de C. Elle est notée C^* . Si \mathbf{l} est tangente à C, C^* vérifie $\mathbf{l}^T C^* \mathbf{l} = 0$*

Théorème A.6.1. *La conique duale d'une conique C peut s'écrire, $C^* = C^{-1}$.*

Définition A.6.2. *Une droite \mathbf{l} (ou un point \mathbf{x}) est en relation polaire à un point \mathbf{x} (ou respectivement à une droite \mathbf{l}) par rapport à une conique C si et seulement si $\mathbf{l} = C\mathbf{x}$ ($\mathbf{l} = C\mathbf{x}$).*

Théorème A.6.2. *Si un point \mathbf{x} est sur une conique C alors la droite polaire du point est tangente à la conique C.*

Si le point est à l'extérieur de la conique, la droite polaire est la droite passant par les deux points d'intersections entre la conique et les tangentes à la conique passant par ce point.

A.7 Transformations projectives et invariants

A.7.1 Invariant projectif

L'invariant projectif le plus connu et le plus utilisé est l'invariant du birapport ou autrement nommé rapport anharmonique. Soit 4 points de \mathbb{P}^1 , toute transformation projective de \mathbb{P}^1 sur lui même conserve le birapport, et ce quelle que soit la représentation en coordonnées homogènes de ces points. Le birapport est défini ci-dessous par :

$$\text{cross}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \frac{|\mathbf{x}_1\mathbf{x}_2||\mathbf{x}_3\mathbf{x}_4|}{|\mathbf{x}_1\mathbf{x}_3||\mathbf{x}_2\mathbf{x}_4|} \text{ avec } |\mathbf{x}_i\mathbf{x}_j| = \det([\mathbf{x}_i \ \mathbf{x}_j]) \quad (\text{A.3})$$

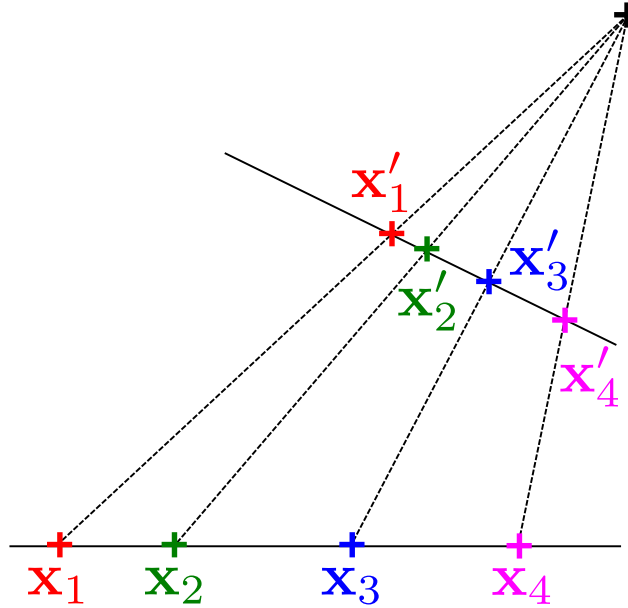


FIGURE A.2 – Projection quelconque de 4 points de \mathbb{P}^1

Sous une transformation projective de \mathbb{P}^2 ou \mathbb{P}^3 qui transforme une droite en une droite, on rappelle qu'une transformation projective conserve l'incidence. Une transformation projective de \mathbb{P}^1 peut être induite de cette transformation en considérant 4 points de la droite ainsi transformée. Sur la FIGURE A.2, nous représentons la perspective d'une droite du plan ce qui induit sur les points des droites respectives une transformation projective de \mathbb{P}^1 . On peut ainsi constater que :

$$\text{cross}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \text{cross}(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \mathbf{x}'_4) \quad (\text{A.4})$$

Cette propriété est intéressante dans le cadre des marqueurs plans, car elle permet ainsi de fournir une contrainte sur la position relative de l'image projective de 4 points quelconques alignés.

A.7.2 Décomposition de l'homographie

Nous rappelons que la transformation projective du plan est généralement décomposée suivant trois types de transformations particulières [30, p.43] :

$$H = H_P H_A H_S = \begin{bmatrix} I_2 & \mathbf{0} \\ \mathbf{v}^T & 1 \end{bmatrix} \begin{bmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

À noter que \mathbf{v} , K , s , R , et \mathbf{t} ne sont pas des paramètres indépendants et doivent se distinguer des paramètres de la caméra. Les trois transformations sont respectivement une transformation perspective, une transformation affine et une similitude.

Une décomposition analogue peut être faite pour les transformations dans l'espace, et il est intéressant de remarquer que la transformation euclidienne ou solide : $\begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ est un cas particulier de similitude où le facteur d'échelle est fixé à 1.

Les paramètres d'une transformation projective peuvent être retrouvés en utilisant des correspondances de points ou de droites, comme nous le verrons dans la section suivante. Il est également possible de simplifier l'estimation en se limitant au calcul de l'une de ses composées grâce à l'estimation de leurs invariants. Les invariants permettent en effet de caractériser certaines transformations sans avoir à calculer les paramètres de toutes les autres. Le cas d'école est le calcul de la ligne de fuite, \mathbf{v} , pour la transformation projective plane qui est l'image par l'homographie de la droite à l'infini \mathbf{l}_∞ . Une fois cette droite estimée il est possible de rectifier l'image affinement, c'est-à-dire de calculer une transformation perspective qui ramène le problème au calcul d'une transformation affine.

A.7.3 Invariants des transformations projectives

Dans cette section, nous souhaitons rappeler des propriétés des invariants des espaces projectifs car ceux-ci seront utiles pour caractériser la nature des transformations que nous étudions par la suite.

La transformation affine a pour propriété de laisser invariante une droite du plan projectif. En effet, nous constatons que la droite ayant pour coordonnées $\mathbf{l}_\infty \sim [0, 0, 1]^T$ reste inchangée. Cette droite est appelée la droite à l'infini, elle contient l'ensemble des points à l'infini du plan projectif. Dans l'espace projectif, son équivalent est le plan à l'infini, noté Π_∞ .

Ces objets géométriques donnent la structure affine de l'espace projectif. Ils sont liés au parallélisme des droites de l'espace, ainsi ils respectent la propriété suivante :

Théorème A.7.1. *Les droites parallèles respectives de l'espace et du plan s'intersectent respectivement sur Π_∞ et \mathbf{l}_∞ .*

Une autre propriété importante, utilisée dans le cas des marqueurs circulaires étudiés dans nos travaux, est la suivante :

Théorème A.7.2. *Le centre d'un cercle (ou respectivement d'une sphère) du plan (ou de l'espace) projectif est en relation polaire avec la droite à l'infini, \mathbf{l}_∞ (ou le plan à l'infini, Π_∞).*

En effet il faut comprendre ici que retrouver l'image de la droite à l'infini ou l'image du centre d'un cercle, noté \mathbf{x}_c , par une transformation projective revient à résoudre le même problème lorsque l'image du cercle, de matrice E , en question est connue. Ainsi nous avons la propriété suivante :

$$E\mathbf{x}_c \sim \mathbf{v}_\infty \tag{A.5}$$

A.8 Rectification affine et rectification métrique

Retrouver les invariants de l'espace projectif permet de retrouver les propriétés familières qui facilitent son étude, parallélisme pour les invariants affines et orthogonalité pour les invariants métriques. Cela permet donc d'effectuer une rectification des espaces projectifs quelconques. Ainsi, nous pouvons calculer les paramètres des transformations qui projettent une primitive observée et son image dans le plan d'une caméra.

Définition A.8.1. *Une rectification affine est une transformation projective qui permet d'obtenir une forme canonique de la droite ou du plan à l'infini. Ainsi, toute transformation du plan projectif qui assure que $\mathbf{l}_\infty \sim [0 \ 0 \ 1]^\top$ est une rectification affine.*

Une similitude de l'espace projectif a pour propriété de laisser invariante une conique particulière de l'espace nommée conique absolue, notée Ω_∞ . La conique absolue est une conique virtuelle dans le sens où aucun point réel ne passe par celle-ci, elle est uniquement composée de points complexes. Les points cycliques, notés \mathbf{I} et \mathbf{J} , sont des points invariants à une similitude du plan projectif. Il s'agit de deux points complexes conjugués qui se trouvent sur la droite à l'infini. Leur nom vient de la propriété suivante qui les caractérise :

Théorème A.8.1. *Les points cycliques passent par une conique non dégénérée si et seulement si cette conique est un cercle.*

Définition A.8.2. *Une rectification métrique correspond à toute transformation projective du plan sur lui même qui permet d'exprimer les points cycliques sous la forme canonique :*

$$\mathbf{I} = [1 \ i \ 0]^\top \text{ et } \mathbf{J} = [1 \ -i \ 0]^\top$$

Avec $i^2 = -1$.

De manière équivalente et sans avoir besoin de distinguer \mathbf{I} et \mathbf{J} , il suffit d'avoir :

$$\mathbf{IJ}^\top + \mathbf{JI}^\top = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

À noter que cette conique dégénérée de rang 2 est nommée conique duale des points cycliques.

A.9 Application au cas d'une projection par une caméra

La projection par une caméra est un cas particulier de projection d'un espace projectif canonique vers un nouvel espace projectif de dimension inférieure, en général \mathbb{P}^3 vers \mathbb{P}^2 .

En supposant un repère quelconque, nous reprenons le modèle de projection à l'équation A.2 de la première section. L'espace projectif de départ est l'espace \mathbb{P}^3 canonique, de sorte que le plan à l'infini et la conique absolue dans cet espace sont :

$$\Pi_\infty \sim [0 \ 0 \ 0 \ 1]^\top \text{ et } \Omega_\infty \sim \mathbf{l}_3 \tag{A.6}$$

Dans le plan image, la conique absolue est projetée en l'image de la conique absolue, notée ω , cette conique particulière aura son importance dans la suite et possède de

nombreuses propriétés. Son image peut être calculée lorsque les paramètres intrinsèques de la caméra sont connus [30, P. 209] :

$$\omega \sim \mathbf{K}^{-\top} \mathbf{K}^{-1} \quad (\text{A.7})$$

Supposons que nous étudions un plan particulier de l'espace projectif, plus précisément le plan passant par l'origine et avec pour normale le vecteur canonique \mathbf{e}_3 . Nous pouvons définir un plan projectif à partir de celui-ci, en prenant pour droite à l'infini $\mathbf{l}_\infty \sim \mathbf{e}_3$. La droite à l'infini de ce plan est matérialisée dans notre espace projectif de dimension supérieure, \mathbb{P}^3 par la droite de l'espace :

$$\mathbf{l}_\infty \sim [0 \ 0 \ 1 \ 1]^\top \quad (\text{A.8})$$

Cette droite est située sur le plan à l'infini de \mathbb{P}^3 et nous pouvons noter qu'elle intersecte la conique absolue en deux points complexes conjugués :

$$\mathbf{I}, \mathbf{J} \sim (1 \ \pm i \ 0)^\top \quad (\text{A.9})$$

Nous reconnaissons bien évidemment les points cycliques du plan projectif. Une représentation de ces objets est donnée ainsi sur la FIGURE A.3.

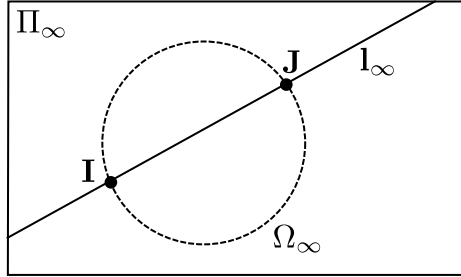


FIGURE A.3 – Relation entre les invariants du plan projectif et de l'espace projectif

L'incidence étant conservée par les transformations projectives, nous rappelons que :

Définition A.9.1. *L'intersection de l'image de la conique absolue et de la ligne de fuite donne l'image des points cycliques.*

Dans la suite, nous garderons ces rappels en tête pour traiter le problème qui nous intéresse, l'estimation de la pose d'une caméra à partir d'un marqueur circulaire. La section suivante propose un état de l'art sur les marqueurs circulaires qui permettra de contextualiser nos travaux. Nous insisterons surtout sur les méthodes de calcul de l'homographie et de la pose qui correspondent au point de départ de nos recherches.

Annexe B

Ajustement d'une ellipse à un ensemble de points

B.1 Formulation du problème

L'ajustement d'une ellipse consiste à trouver l'ellipse la plus proche en distance euclidienne de points donnés. L'ellipse peut être paramétrée par 6 paramètres condensés dans un vecteur homogène $\theta = (a, b, c, d, e, f)$ de sorte qu'un point sur l'ellipse de vecteur $\mathbf{x} = [x \ y \ 1]^\top$ respecte :

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (\text{B.1})$$

Les coefficients de θ paramètrent la forme matricielle de la conique :

$$\mathbf{E} = \begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{f}{2} \\ \frac{d}{2} & \frac{f}{2} & f \end{bmatrix} \quad (\text{B.2})$$

On ajoute au problème une contrainte pour vérifier que la conique soit bien une ellipse et non pas une conique quelconque.

$$\Delta = 4ac - b^2 > 0 \quad (\text{B.3})$$

Le problème principal est que la distance euclidienne entre un point et une ellipse ne peut pas être obtenue directement et nécessite la résolution d'un système polynomial de degré 4 (équation quartique). Ainsi le problème suivant ne peut pas être résolu directement :

$$\min_{\theta} \left\{ \sum_{i=1}^n d(\mathbf{x}_i, \theta) \right\} \text{ avec } \mathbf{x}_i \in \mathbb{R}^2, i \in \mathbb{N} \quad (\text{B.4})$$

Une solution pour résoudre ce problème est de transformer l'espace des données dans un espace généralement de dimension plus grande où la mesure de distance peut être formulée linéairement. Pour la distance à une conique, on peut utiliser l'espace de dimension 6 tel que les données deviennent $\epsilon_i = [x_i^2 \ x_i y_i \ y_i^2 \ x_i \ y_i \ 1]^\top \in \mathbb{R}^6$.

Le problème revient alors à trouver un vecteur de paramètre $\theta \in \mathbb{R}^6$ tel que $\forall i, \epsilon_i^\top \theta = 0$. L'ajout d'une contrainte quadratique permet de garantir que la conique ainsi estimée est bien une ellipse et pas une conique quelconque, puisqu'on travaille en coordonnées homogènes on peut fixer Δ , tel que $\Delta = 4ac - b^2 = 1$. La méthode au moindre carré linéaire vise donc à trouver le paramètre θ en minimisant l'expression sous la forme :

$$\|\mathbf{A}\theta\|^2 = \theta^\top \mathbf{A}^\top \mathbf{A} \theta \quad (\text{B.5})$$

avec la contrainte :

$$\theta^\top \mathbf{C} \theta = \text{constante} > 0 \quad (\text{B.6})$$

où \mathbf{A} a pour lignes l'ensemble des vecteurs ϵ_i^\top et :

$$\mathbf{C} = \begin{bmatrix} \mathbf{M} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{bmatrix} \text{ avec } \mathbf{M} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix} \quad (\text{B.7})$$

Cette première méthode d'estimation de l'ellipse est très utilisée en raison de sa simplicité et de son efficacité elle a été proposée par Fitzgibbon et Halir [21, 28]. En utilisant le multiplicateur de Lagrange on cherche à minimiser :

$$L(\theta, \lambda) = \theta^\top \mathbf{A}^\top \mathbf{A} \theta - \lambda \theta^\top \mathbf{C} \theta + \lambda \quad (\text{B.8})$$

On a alors les conditions d'optimalité données par :

$$\frac{\partial L}{\partial \theta} = \theta^\top A^\top A - \lambda \theta^\top C = 0 \quad (\text{B.9})$$

$$\frac{\partial L}{\partial \lambda} = \theta^\top C \theta - 1 = 0 \quad (\text{B.10})$$

La première condition d'optimalité nous donne ainsi :

$$A^\top A \theta = \lambda C \theta \quad (\text{B.11})$$

Le vecteur θ optimal est donc un vecteur propre généralisé de $(A^\top A, C)$ associé à la valeur propre λ . À noter que l'on a donc également :

$$\|A\theta\|^2 = \theta^\top A^\top A \theta = \theta^\top \lambda C \theta = \lambda \quad (\text{B.12})$$

Ainsi pour résoudre le problème il faut prendre le vecteur propre de (AA^\top, C) associé à la plus petite valeur propre λ . Cette méthode n'est cependant pas parfaite, la transformation de l'espace des données induit un biais important sur l'ellipse estimée, les méthodes présentées dans la suite ont pour but de réduire ce biais ou de trouver une meilleure mesure de distance.

B.2 Résolution avec des distances alternatives

Pour simplifier la résolution du problème d'ajustement (B.4), il est possible de remplacer la distance euclidienne par une distance approchée afin de pouvoir résoudre le problème au moindre carré.

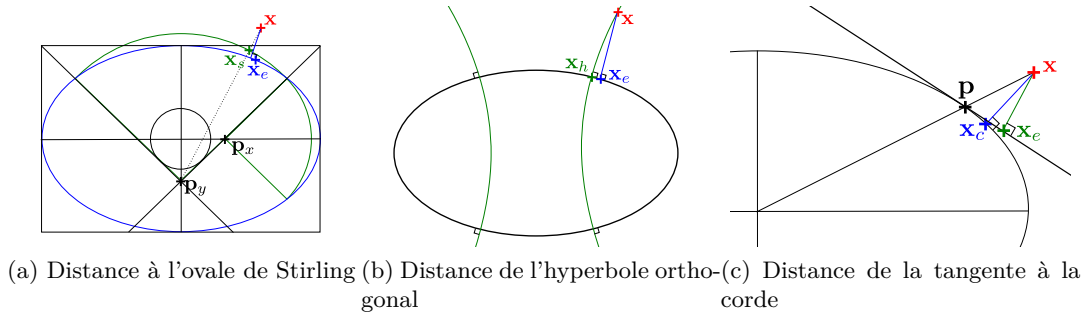


FIGURE B.1 – Différentes mesures de distances d'un point à une ellipse

Un premier exemple de cette approximation est l'ovale de Stirling, cette distance se mesure en construisant une approximation de l'ellipse par 4 arcs de cercle. La FIGURE B.1a illustre leurs constructions, un rectangle englobant est d'abord construit autour de l'ellipse avec un cercle interne. Des tangentes au cercle interne et passant par les quatre coins du rectangle sont ensuite utilisées. Les centres des arcs p_x et p_y sont donnés par l'intersection de ces tangentes avec les deux axes principaux de l'ellipse. Les arcs peuvent finalement être construits avec leurs centres et séparés en 4 arcs par les tangentes comme illustré sur la figure. La distance approchée est alors la distance entre le point, x , et l'arc de cercle contenu dans le même quadrant. Cette construction mène à une distance simplement quadratique relativement aux coordonnées du point x mais rajoute une condition sur le test du quadrant [70].

Rosin compare cette distance à l'hyperbole orthogonale [70], il montre que cette dernière est plus précise et donne une meilleure estimation pour l'ellipse. Cette distance illustrée sur la FIGURE B.1b se base sur le fait que les hyperboles confocales¹ intersectent toujours l'ellipse orthogonalement. Ainsi la distance proposée est celle entre le point \mathbf{x} et le point, \mathbf{x}_h , intersection de l'ellipse avec l'hyperbole confocale passant par le point, \mathbf{x} . La mesure de cette distance consiste à calculer les 4 points d'intersection puis à mesurer la plus courte distance entre ces points et le point \mathbf{x} . Elle est donc également quadratique et avec une condition.

Plus récemment Ou *et al.* [62] présentent la distance de la corde tangente, celle-ci consiste à construire une tangente de l'ellipse au point \mathbf{p} , point d'intersection entre l'ellipse et son rayon passant par le point \mathbf{x} . La distance de la corde tangente est alors la distance entre le point \mathbf{x} et cette tangente comme illustré sur la FIGURE B.1c.

B.3 Ajustement d'une ellipse au maximum de vraisemblance

La méthode du maximum de vraisemblance propose également de supprimer le biais de la résolution géométrique en utilisant les matrices de covariances associées à l'incertitude sur les points de l'ellipse afin d'estimer une distance de *Mahalanobis* des points réels $\hat{\mathbf{X}}_i = [\hat{x}_i \ \hat{y}_i]^\top$ sur l'ellipse à leurs observations $\mathbf{X}_i = [x_i \ y_i]^\top \in \mathbb{R}^2$. Cette distance se mesure ainsi :

$$\frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{X}}_i - \mathbf{X}_i)^\top \Sigma_{\mathbf{X}_i}^{-1} (\hat{\mathbf{X}}_i - \mathbf{X}_i) \quad (\text{B.13})$$

où $\Sigma_{\mathbf{X}_i}$ est la matrice de covariance associée à l'incertitude du point de coordonnées \mathbf{X}_i et avec les contraintes :

$$[\hat{x}_i \ \hat{y}_i \ 1] \mathbf{E}(\theta) [\hat{x}_i \ \hat{y}_i \ 1]^\top \quad (\text{B.14})$$

La minimisation de cette distance a l'avantage de pondérer les points selon l'incertitude de leur position dans l'image. La résolution de ce problème n'est cependant pas triviale puisque les paramètres de l'ellipse condensés dans le vecteur θ n'interviennent pas directement dans le produit à minimiser mais uniquement dans la contrainte ($\mathbf{E}(\theta)$ est l'ellipse sous forme matricielle de l'expression B.2).

Une approche pour résoudre ce problème consiste à le résoudre dans l'espace reparamétrisé des données (\mathbb{R}^5), c'est-à-dire à partir des vecteurs ϵ_i . Si l'hypothèse d'un bruit gaussien pour la position des points de coordonnées \mathbf{X}_i dans le plan est envisageable ce n'est pas directement possible pour les données dans l'espace reparamétrisé de coordonnées ϵ_i , en faisant cependant cette approximation, raisonnable pour un bruit de faible variance, il est possible d'obtenir une mesure qui peut être directement minimisée.

Pour cela on calcule les matrices de covariance des données dans cet espace, Σ_{ϵ_i} . On obtient alors l'équation :

$$\frac{1}{N} \sum_{i=1}^N (\hat{\epsilon}_i - \mathbf{X}_i)^\top \Sigma_{\epsilon_i}^{-1} (\hat{\epsilon}_i - \epsilon_i) \quad (\text{B.15})$$

avec les contraintes linéaires :

$$\hat{\epsilon}_i^\top \theta = 0 \quad (\text{B.16})$$

1. Des coniques confocales partagent les mêmes foyers

et grâce au multiplicateur de Lagrange on obtient une fonction à minimiser sans contrainte qui dépend cette fois de θ :

$$\frac{1}{N} \sum_{i=1}^N \frac{\epsilon_i^\top \theta \theta^\top \epsilon_i}{\theta^\top \Sigma_{\epsilon_i} \theta} \quad (\text{B.17})$$

En fait il est possible d'approximer Σ_{ϵ_i} à l'ordre 1 en utilisant la jacobienne de $\frac{\partial \epsilon_i}{\partial \mathbf{x}_i}$, ce qui donne ainsi :

$$\frac{1}{N} \sum_{i=1}^N \frac{\epsilon_i^\top \theta \theta^\top \epsilon_i}{\theta^\top \frac{\partial \epsilon_i}{\partial \mathbf{x}_i} \Sigma_{\mathbf{x}_i} \frac{\partial \epsilon_i}{\partial \mathbf{x}_i}^\top \theta} \quad (\text{B.18})$$

On obtient alors l'erreur de Sampson, qu'on peut interpréter comme l'erreur avec une distance algébrique ajustée par des poids réduisant le biais d'estimation dû à la distance algébrique. Chojnacki *et al.* [12] proposent un schéma itératif de minimisation de l'erreur de Sampson pour des coniques, cette minimisation cependant ne prend pas en compte la spécificité des coniques dont on fait l'estimation, comme l'ellipse dans notre cas.

Le problème de la minimisation de l'expression B.18 est qu'on ne prend pas en compte le caractère elliptique de la conique. Szpak *et al.* [77] proposent de reparamétriser la conique de sorte que celle-ci reste dans l'espace des ellipses c'est-à-dire tel que $\Delta > 0$. Il propose ainsi d'écrire :

$$\theta = [1 \quad 2p \quad p^2 + q^{-2} \quad r \quad s \quad t]^\top \quad (\text{B.19})$$

avec un nouveau vecteur de paramètre $\eta = [p \quad q \quad r \quad s \quad t]^\top$. Les auteurs adaptent ensuite un algorithme d'optimisation de l'expression B.18 basé sur Levenberg-Marquardt. Finalement les auteurs proposent de calculer la covariance associée aux paramètres géométriques de l'ellipse de leur estimation, ce qui permet de mesurer une région de confiance pour l'ellipse obtenue. Nous avons donc retenu cet algorithme à la fois pour sa précision d'estimation des paramètres de l'ellipse, mais aussi pour la mesure de l'incertitude que nous utilisons dans le chapitre 3.

Annexe C

Résultats expérimentaux supplémentaires

C.1 Description du protocole d'évaluation

L'objectif des tests présentés dans cette section est d'évaluer la méthode proposée pour estimer la pose d'un cercle. Nous avons effectué ces tests sur des images synthétiques et réelles dans les conditions illustrées sur la FIGURE C.1.

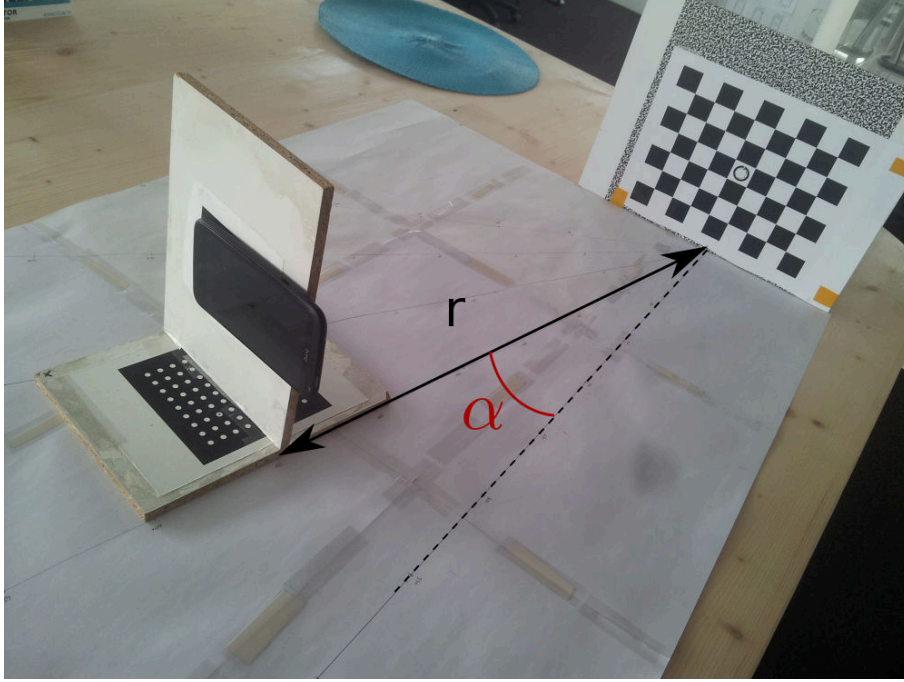


FIGURE C.1 – Photo de notre installation

Afin de ne pas tester toutes les poses, nous avons simplifié le problème pour limiter les échantillons à un nombre plus restreint de paramètres.

Tout d'abord, nous supposons que la caméra vise le centre du marqueur, c'est-à-dire que l'axe principal de la caméra passe par le centre du marqueur, voir la FIGURE C.2. En pratique, l'axe principal passe par un point du plan très proche du centre mais à plus d'un rayon de cercle, soit $\frac{D}{2}$, du centre du marqueur. Ainsi, la distance entre ce point est bien inférieure à r la distance entre le marqueur et la caméra.

L'angle γ correspond à l'angle de rotation de la caméra par rapport à son axe principal. Si l'on suppose que l'axe principal est orthogonal au plan image, l'angle γ n'aura pas d'autre effet sur la projection qu'une rotation du plan image sur lui même. Dans nos expériences, nous le supposons donc fixe car sa variation n'ajoute pas de nouvelles configurations.

De même l'angle β correspond à l'angle de rotation du marqueur autour de lui même, celui-ci n'apporte pas de nouvelles configurations du cercle par rapport à la caméra, on le suppose donc fixe également.

Il nous reste donc deux variables que nous faisons varier pour évaluer différentes configurations de pose :

- L'angle α d'inclinaison de l'axe principal par rapport au plan de support ;
- La distance de la caméra au centre du marqueur, notée r .

Il est important de noter que nous retrouvons la condition de la section 4.3.3 du chapitre 4 et ainsi l'ambiguïté peut être levée tant que nous continuons à respecter ces conditions.

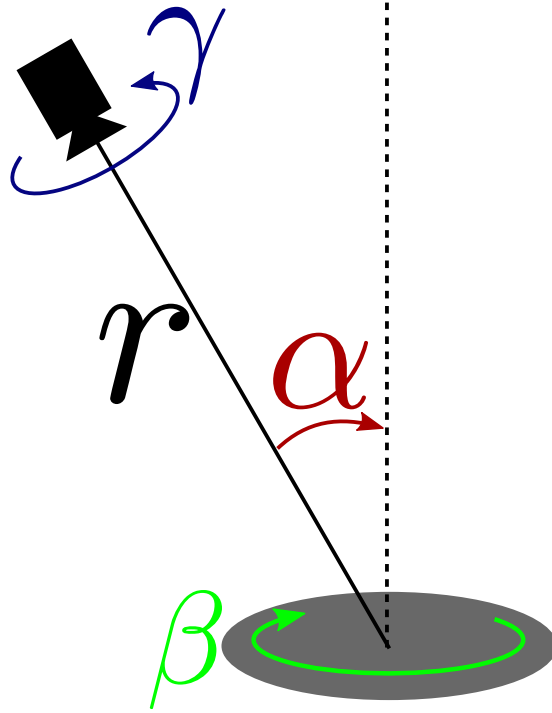


FIGURE C.2 – Paramétrage de la pose du cercle par rapport à la caméra.

Nous savons que l'introduction des paramètres par défaut pour la caméra, comme proposée dans le chapitre 3, a un impact négatif sur la précision de l'estimation de la pose. Par conséquent, l'un des objectifs de cette expérience est d'évaluer la sensibilité de la méthode proposée à l'imprécision du paramètre focal de la caméra. L'observation de la distribution de la distance focale de diverses caméras de smartphones, voir section 3.5.1 du chapitre 3, révèle que tous les équivalents focaux 35 *mm* sont inclus dans un intervalle $[-30\%, +30\%]$ autour de la valeur moyenne. Après calibrage de la caméra utilisée dans l'expérience, nous avons réalisé l'expérience en utilisant 5 valeurs de distance focale qui couvre cet intervalle. Ainsi, la distance focale réelle a été multipliée par les facteurs, $\{0.7, 0.85, 1.0, 1.15, 1.3\}$ pour voir l'impact d'une erreur de distance focale sur les résultats.

Afin de générer des images synthétiques, nous avons simulé une caméra synthétique en utilisant les mêmes paramètres que la caméra réelle utilisée dans nos tests avec une résolution de 1280×720 pixels. L'image synthétique a été échantillonnée avec une résolution plus élevée puis sous-échantillonnée pour éviter d'avoir de l'*aliasing* car ce phénomène est problématique, notamment pour la détection du contour.

Les images réelles ont été obtenues avec la caméra d'un smartphone (HTC One X+) qui a été calibrée avec la bibliothèque *OpenCV*¹.

Dans les deux cas, nous supposons les ellipses correctement détectées dans l'image. Les points du contour ont été détectés par la méthode présentée dans le chapitre 1 et les ellipses sont estimées avec la méthode de Szpak et al. [77].

Enfin, nous évaluons la qualité des résultats obtenus en utilisant trois mesures d'erreur différentes relatives à la pose et à la précision de la reprojection : l'erreur sur l'estimation de la normale du plan de support, l'erreur sur l'estimation de la position du

1. <https://opencv.org/>

marqueur et l'erreur de reprojection de points 3D proches du marqueur dans l'image.

Chaque courbe illustre les résultats obtenus en modifiant la distance focale utilisée pour l'estimation de la pose. Les erreurs résultantes sont affichées en fonction de la distance r dans l'intervalle $[15 \times D, 45 \times D]$ où D est le diamètre du marqueur. Cet intervalle est lié aux distances utilisées pour être capable de détecter et de reconnaître un marqueur dans une application de réalité augmentée. Cette distance dépend de la résolution de la caméra, il faut notamment que le marqueur occupe au moins 80 pixels dans l'image. Nous présentons également les résultats pour trois valeurs différentes de l'angle α , soit $\alpha \in \{15^\circ, 30^\circ, 45^\circ\}$, affichées sous formes de trois sous graphes.

C.2 Analyse de nos résultats

Les résultats sur les images synthétiques sont présentés sur la FIGURE C.3. Nous montrons l'erreur d'estimation de l'orientation du plan sur la FIGURE C.3a. Nous pouvons voir que dans le cas calibré (courbe noire), l'erreur augmente légèrement avec la distance mais diminue avec l'angle. Cette observation est en fait intuitive : plus le cercle est éloigné, plus nous perdons de la précision sur l'ellipse détectée et il en va de même pour la normale calculée à partir de celle-ci. Lorsque le cercle est observé de face (figure de gauche avec un angle plus faible), l'ellipse devient plus circulaire et la précision sur la normale calculée diminue à mesure que la solution s'approche d'un cas dégénéré. Ces deux observations sont vraies pour l'erreur de reprojection dans la FIGURE C.3c et pour les images réelles dans la FIGURE C.4 également.

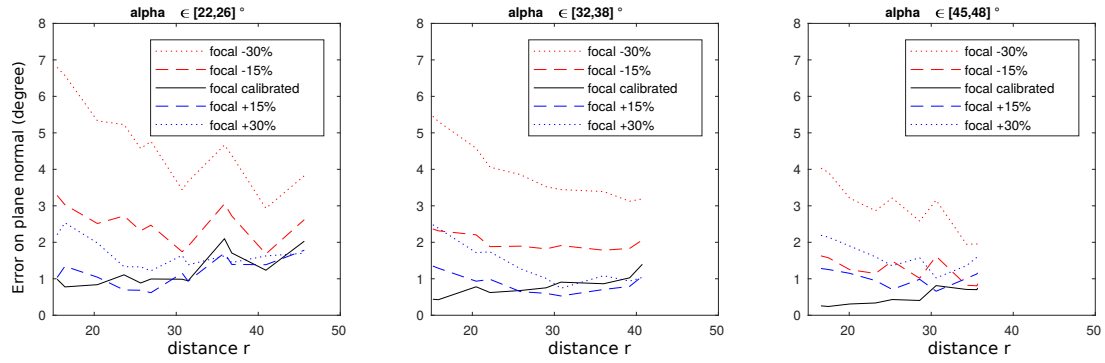
Le résultat que nous voulons souligner est que l'erreur sur la normale estimée augmente en fait assez peu malgré la modification appliquée sur la distance focale, c'est-à-dire que les courbes rouges et bleues de la FIGURE C.3a restent assez proches de la courbe noire. Cela signifie qu'un calibrage imprécis sur la caméra n'affecte pas autant qu'on pourrait le penser le calcul de la normale. De plus, nous pouvons voir que l'erreur relative du cas avec modification (courbe rouge ou bleue) par rapport à l'erreur du cas calibré (courbe noire) diminue avec la distance. Cette observation est encore plus remarquable sur l'erreur de reprojection dans la FIGURE C.3c.

Dans la FIGURE C.3b où le résultat sur l'erreur de position est affiché, on peut voir que dans le cas calibré la précision en position reste faible et ne dépend ni de la distance à la caméra ni de l'angle, α , entre le plan de support et la caméra. Dans les cas non calibrés, comme prévu, la détection des ellipses devient moins précise lorsque la distance augmente et, par conséquent, la qualité de l'estimation de la position du marqueur est également fortement affectée.

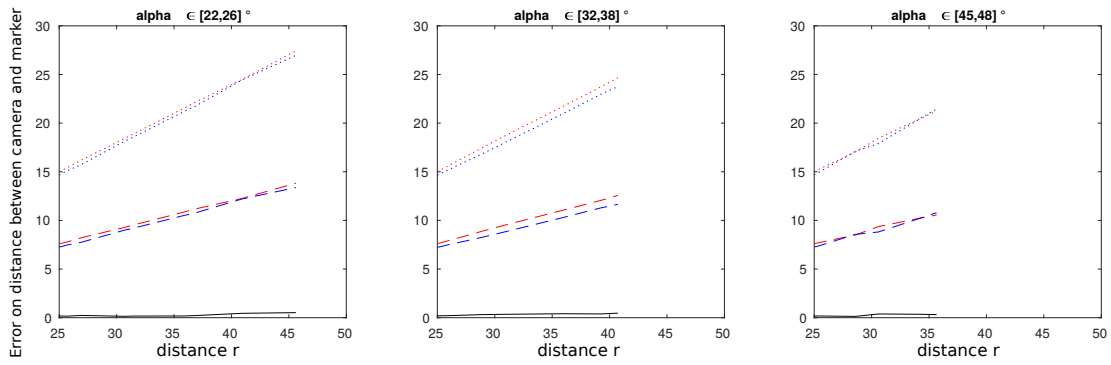
En fait, l'erreur de position augmente linéairement lorsque la distance augmente, voir la FIGURE C.3b. Cette observation correspond à l'effet d'un *travelling* contrarié appliqué au marqueur. C'est le véritable effet négatif de l'utilisation d'une mauvaise distance focale. Une variation de la distance focale utilisée pour le calcul de la pose du marqueur provoque une variation de la profondeur estimée du marqueur à partir de l'image.

Cependant, l'erreur de reprojection présentée dans la FIGURE C.3c reste faible lorsque la distance augmente. Cela signifie que l'utilisation de paramètres génériques n'affecte pas la qualité de la reprojection dans un contexte où le marqueur est loin de la caméra et les points reprojétés proches du marqueur. Ce résultat montre que cette approche est adaptée au contexte de réalité augmentée.

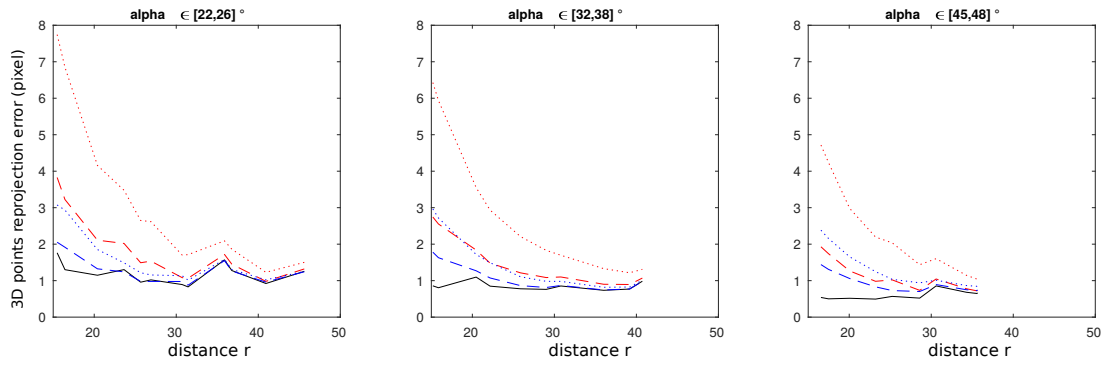
La FIGURE C.4 nous permet de présenter des observations et des conclusions similaires sur des images réelles. La principale différence est que la précision de l'estimation



(a) Erreur d'orientation de la normale



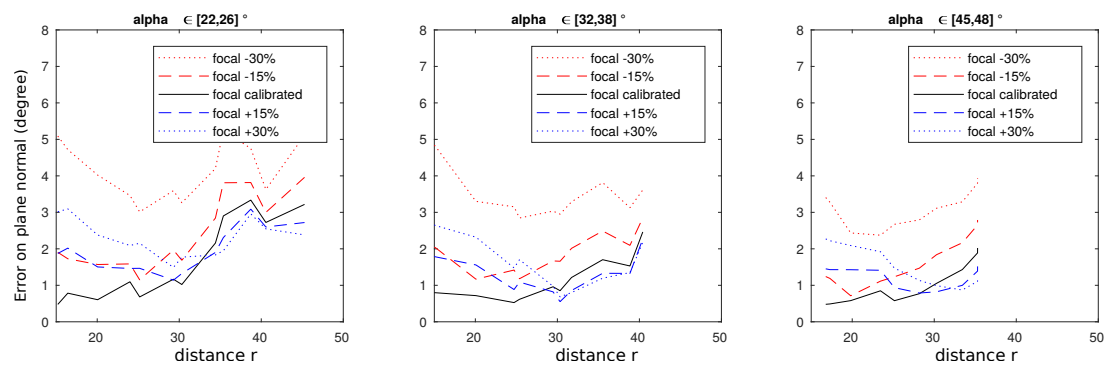
(b) Erreur de position du marqueur



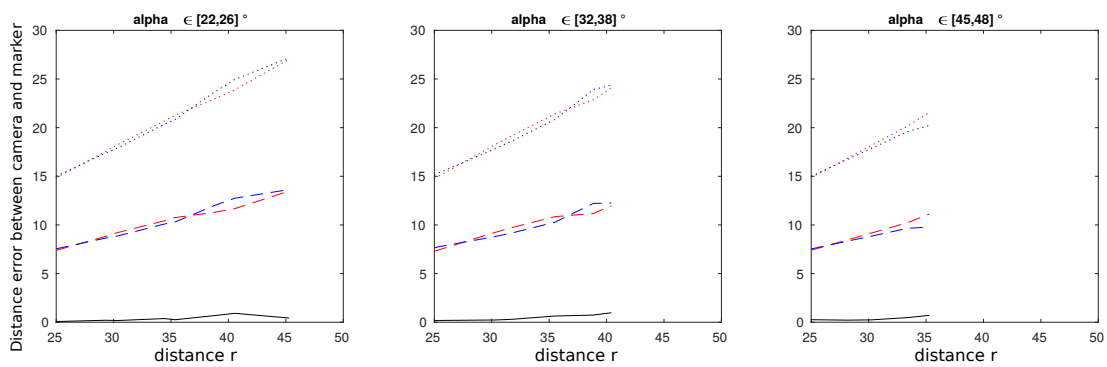
(c) Erreur de reprojection de points 3D

FIGURE C.3 – Résultats sur les images synthétiques.

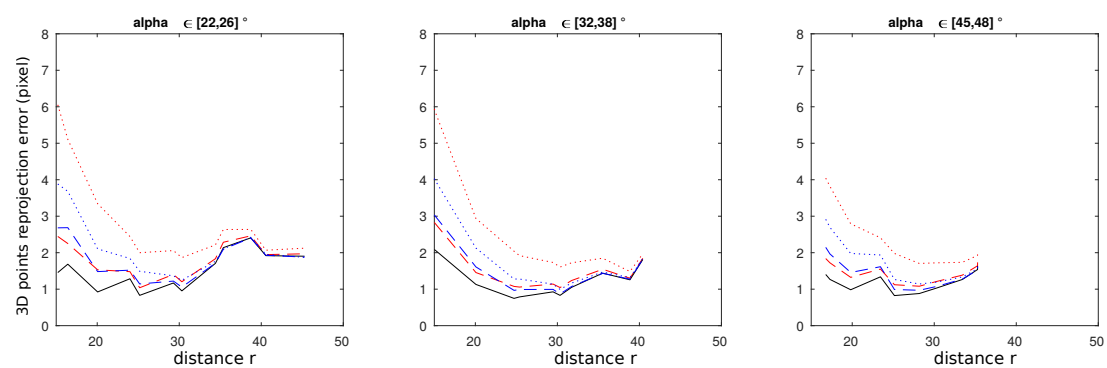
semble être plus sensible à la distance et donc à la taille de l'ellipse sur l'image. Cela est cohérent avec le fait que la détection de l'ellipse est plus difficile sur les images réelles.



(a) Erreur d'orientation de la normale



(b) Erreur de position du marqueur



(c) Erreur de reprojection de points 3D

FIGURE C.4 – Résultats sur les images réelles

Annexe D

Preuve en calcul formel *Matlab*

D.1 Code formel pour la preuve de la proposition 4.2.6

```
1 syms x y R theta real
2 Q_0 = [1,0,-x;0,1,-y;-x,-y,x^2+y^2-R^2];
3 E = [1,0,0;0,1,cos(theta);0,0,-sin(theta)];
4 Q = simplify(sin(theta)^2*transpose(inv(E))*Q_0*inv(E));
5 [V,D] = eig(Q);
6 tau = 4*y^3*cos(theta) + 2*x^2*y^2 - 2*R^2*cos(2*theta) + ...
7     2*y^2*cos(2*theta) + 4*y*cos(theta) + R^4 + 2*x^2 + x^4 + ...
8     4*y^2 + y^4 - 2*R^2*x^2 - 2*R^2*y^2 - ...
9     4*R^2*y*cos(theta) + 4*x^2*y*cos(theta) + 1;
10 lambda_1 = y*cos(theta) - R^2/2 - sqrt(tau)/2 + x^2/2 + y^2/2 + 1/2;
11 lambda_2 = sin(theta)^2;
12 lambda_3 = y*cos(theta) - R^2/2 + sqrt(tau)/2 + x^2/2 + y^2/2 + 1/2;
13 % We write D2 the degenerate conic we search
14 % and check that its signature is (1,1)
15 D2 = Q - lambda_2*eye(3);
16 D2_eigs = simplify(eig(D2),10);
17 IND1 = find(D2_eigs==0);
18 IND23 = find(D2_eigs~=0);
19 check_sign = simplify(prod(D2_eigs(IND23)),10);
20 % Product of non null eigenvalues
21 % We can find a more simple expression
22 check_sign_0 = -sin(theta)^2*(cos(theta)+y)^2 + x^2);
23 simplify(check_sign_0-check_sign,10) % ans = 0
24
25 % We can check the two line composing this conic
26 l_0 = [D2(1,3);D2(2,3);D2(3,3)/2];
27 l_inf = [0;0;1];
28 simplify(l_0*transpose(l_inf)+l_inf*transpose(l_0) == D2,10)
29 % ans =
30 % [ TRUE, TRUE, TRUE]
31 % [ TRUE, TRUE, TRUE]
32 % [ TRUE, TRUE, TRUE]
33 %
```

```

34 % Now we want to find the points x_i
35 % corresponding to the generalized eigenvectors
36 x1 = simplify(V(:,2),10);
37 x3 = simplify(V(:,3),10);
38 % Simplification of x1 with tau
39 syms tau
40 x1_1 = -(x*(cos(2*theta) + sqrt(tau) + 2*y*cos(theta) - R^2 + ...
41   x^2 + y^2))/(2*sin(theta)*(cos(theta)^2 + ...
42   2*y*cos(theta) + x^2 + y^2));
43 x1_2 = -(2*y + cos(3*theta) + cos(theta) - 2*R^2*cos(theta) + ...
44   2*x^2*cos(theta) + 4*y*cos(2*theta) + 6*y^2*cos(theta) + ...
45   2*y*sqrt(tau) + 2*cos(theta)*sqrt(tau) - 2*R^2*y + 2*x^2*y + ...
46   2*y^3)/(4*sin(theta)*(cos(theta)^2 + ...
47   2*y*cos(theta) + x^2 + y^2));
48 simplify(x1 - [x1_1;x1_2;1],10) % ans = 0
49 % Simplification of x3 with tau
50 x3_1 = -(x*(cos(2*theta) - sqrt(tau) + 2*y*cos(theta) - R^2 + ...
51   x^2 + y^2))/(2*sin(theta)*(cos(theta)^2 + 2*y*cos(theta) + ...
52   x^2 + y^2));
53 x3_2 = -(2*y + cos(3*theta) + cos(theta) - 2*R^2*cos(theta) + ...
54   2*x^2*cos(theta) + 4*y*cos(2*theta) + 6*y^2*cos(theta) - ...
55   2*y*sqrt(tau) - 2*cos(theta)*sqrt(tau) - 2*R^2*y + 2*x^2*y + ...
56   2*y^3)/(4*sin(theta)*(cos(theta)^2 + 2*y*cos(theta) + x^2 + y^2));
57 simplify(x3 - [x3_1;x3_2;1],10) % ans = 0
58 % Check if x1 and x3 lie on the same side of l_0 and l_inf
59 simplify(transpose(l_inf)*[x1_1;x1_2;1],10) % ans = 1
60 simplify(transpose(l_inf)*[x3_1;x3_2;1],10) % ans = 1
61 % We can note that tau is positive so we can easily
62 % determine if l_0 separates x1 and x3
63 simplify(transpose(l_0)*[x1_1;x1_2;1],10) % ans = -tau^(1/2)/2 < 0
64 simplify(transpose(l_0)*[x3_1;x3_2;1],10) % ans = tau^(1/2)/2 > 0

```

D.2 Code formel pour la preuve de la proposition 4.3.1

```

1 % Complete the unknown rotation with a parameter a
2 syms a
3 r1_0 = [1;0;0];
4 r2_0 = [0;sin(theta);cos(theta)];
5 r1 = cos(a)*r1_0 - sin(a)*r2_0;
6 r2 = sin(a)*r1_0 + cos(a)*r2_0;
7 r3 = [0;cos(theta);-sin(theta)];
8 O = [r1,r2,r3];
9 H = transpose(O)*E;
10
11 % Now we use the back projection of the line at infinity
12 % of the image plane into the new plane to
13 % test if points lie both on the same side of the camera
14 l_inf_back = simplify(transpose(inv(E))*transpose(H)*l_inf);

```

```

15
16 s_1 = simplify(transpose(l_inf_back)*[x1_1;x1_2;1],10);
17 s_3 = simplify(transpose(l_inf_back)*[x3_1;x3_2;1],10);
18 tau_0 = 4*y^3*cos(theta) + 2*x^2*y^2 - 2*R^2*cos(2*theta) + ...
19 2*y^2*cos(2*theta) + 4*y*cos(theta) + R^4 + 2*x^2 + x^4 + ...
20 4*y^2 + y^4 - 2*R^2*x^2 - 2*R^2*y^2 - ...
21 4*R^2*y*cos(theta) + 4*x^2*y*cos(theta) + 1;
22 test = simplify(subs(s_1*s_3,tau,tau_0),10);
23 % The denominator is postive as long as |x|>0 or |cos(t)+y|>0
24 % and theta > 0
25 denominator = ((cos(theta)+y)^2 + x^2);
26 test_0 = simplify(test*denominator,10)
27 test_1 = (x^2+y^2+cos(theta)*(cos(theta)*(y^2-R^2) + ...
28 y*(x^2+y^2-R^2+1)))/denominator;
29 simplify(test_1-test,10) % ans = 0

```


Bibliographie

- [1] Anubhav AGARWAL, CV JAWAHAR et PJ NARAYANAN : A survey of planar homography estimation techniques. Centre for Visual Information Technology, Tech. Rep. IIIT/TR/2005/12, 2005.
- [2] Luis ALVAREZ et Vicent CASELLES : Homography estimation using one ellipse correspondence and minimal additional information. In IEEE International Conference on Image Processing (ICIP), 2014.
- [3] Alessandro BENINI, Matthew J. RUTHERFORD et Kimon P. VALAVANIS : Real-time, gpu-based pose estimation of a uav for autonomous takeoff and landing. In IEEE International Conference on Robotics and Automation (ICRA), pages 3463–3470, 2016.
- [4] Filippo BERGAMASCO, Andrea ALBARELLI, Luca COSMO, Emanuele RODOLA et Andrea TORSELLO : An accurate and robust artificial marker based on cyclic codes. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 38(12):2359–2373, 2016.
- [5] Filippo BERGAMASCO, Andrea ALBARELLI, Emanuele RODOLA et Andrea TORSELLO : Rune-tag : A high accuracy fiducial marker with strong occlusion resilience. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), pages 113–120, 2011.
- [6] Tolga BIRDAL, Ievgeniia DOBRYDEN et Slobodan ILIC : X-tag : A fiducial tag for flexible and accurate bundle adjustment. In IEEE 4th International Conference on 3D Vision (3DV), pages 556–564, 2016.
- [7] Lilian CALVET, Pierre GURDJOS, Carsten GRIWODZ et Simone GASPARINI : Detection and accurate localization of circular fiducials under highly challenging conditions. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), 2016.
- [8] Federico CAMPOSECO, Andrea COHEN, Marc POLLEFEYS et Torsten SATTLER : Hybrid camera pose estimation. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [9] J. CANNY : A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 8(6):679–698, 1986.
- [10] Qian CHEN, Haiyuan WU et Toshikazu WADA : Camera calibration with two arbitrary coplanar circles. In European Conference on Computer Vision (ECCV), pages 521–532. Springer, 2004.
- [11] Qian CHEN, Haiyuan WU et Toshikazu WADA : Camera calibration with two arbitrary coplanar circles. In European Conference on Computer Vision (ECCV), pages 521–532. Springer, 2004.

- [12] Wojciech CHOJNACKI, Michael J. BROOKS, Anton VAN DEN HENGEL et Darren GAWLEY : On the fitting of surfaces to data with covariances. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 22(11):1294–1303, 2000.
- [13] Toby COLLINS et Adrien BARTOLI : Infinitesimal plane-based pose estimation. International journal of computer vision, 109(3):252–286, 2014.
- [14] N. DALAL et B. TRIGGS : Histograms of oriented gradients for human detection. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), pages 886–893, 2005.
- [15] Joseph DEGOL, Timothy BRETLE et Derek HOIEM : Chromatag : a colored marker and fast detection algorithm. In IEEE International Conference on Computer Vision (ICCV), pages 1472–1481, 2017.
- [16] Joseph DEGOL, Timothy BRETLE et Derek HOIEM : Improved structure from motion using fiducial marker matching. In European Conference on Computer Vision (ECCV), pages 273–288, 2018.
- [17] Matthew J. EDWARDS, Michael P. HAYES et Richard D. GREEN : High-accuracy fiducial markers for ground truth. In IEEE International Conference on Image and Vision Computing New Zealand (IVCNZ), pages 1–6, 2016.
- [18] Georgios D. EVANGELIDIS et Emmanouil Z. PSARAKIS : Parametric image alignment using enhanced correlation coefficient maximization. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 30(10):1858–1865, 2008.
- [19] Mark FIALA : Artag, a fiducial marker system using digital techniques. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), volume 2, pages 590–596. IEEE, 2005.
- [20] Mark FIALA : Designing highly reliable fiducial markers. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 32(7), 2010.
- [21] Andrew FITZGIBBON, Maurizio PILU et Robert B. FISHER : Direct Least Squares Fitting of Ellipses. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 1996.
- [22] Michele FORNACIARI et Andrea PRATI : Very fast ellipse detection for embedded vision applications. In IEEE 6th International Conference on Distributed Smart Cameras (ICDSC), pages 1–6, 2012.
- [23] Sergio GARRIDO-JURADO, Rafael MUNOZ-SALINAS, Francisco José MADRID-CUEVAS et Rafael MEDINA-CARNICER : Generation of fiducial marker dictionaries using mixed integer linear programming. Pattern Recognition, 51:481–491, 2016.
- [24] Shatadal GHOSH, Ranjit RAY, Siva Ram Krishna VADALI, Sankar Nath SHOME et Sambhunath NANDY : Reliable pose estimation of underwater dock using single camera : a scene invariant approach. Machine Vision and Applications, 27(2):221–236, 2016.
- [25] A Ardeshir GOSHTASBY : Image registration : Principles, tools and methods. Springer Science & Business Media, 2012.
- [26] Pierre GURDJOS, Jun-Sik KIM et In-So KWEON : Euclidean structure from confocal conics : theory and application to camera calibration. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), volume 1, pages 1214–12221, 2006.
- [27] Pierre GURDJOS, Peter STURM et Yihong WU : Euclidean structure from $n \geq 2$ parallel circles : theory and algorithms. In European Conference on Computer Vision (ECCV), pages 238–252. Springer, 2006.

-
- [28] Radim HALIR et Jan FLUSSER : Numerically stable direct least squares fitting of ellipses. In Citeseer 6th International Conference in Central Europe on Computer Graphics and Visualization (WSCG), volume 98, pages 125–132, 1998.
- [29] Christopher G HARRIS, Mike STEPHENS et al. : A combined corner and edge detector. In Alvey vision conference, pages 10–5244. Citeseer, 1988.
- [30] Richard HARTLEY et Andrew ZISSERMAN : Multiple view geometry in computer vision. Cambridge university press, 2003.
- [31] M HASSABALLAH, Aly Amin ABDELMGEID et Hammam A ALSHAZLY : Image features detection, description and matching. In Image Feature Detectors and Descriptors, pages 11–45. Springer, 2016.
- [32] Danying HU, Daniel DETONE et Tomasz MALISIEWICZ : Deep charuco : Dark charuco marker pose estimation. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 8436–8444, 2019.
- [33] Bin HUANG, Yongrong SUN et Qinghua ZENG : General fusion frame of circles and points in vision pose estimation. Optik, 154, 2018.
- [34] Bin HUANG, Yongrong SUN, Yunfeng ZHU, Zhi XIONG et Jianye LIU : Vision pose estimation from planar dual circles in a single image. Optik, 127(10):4275–4280, 2016.
- [35] Haifei HUANG, Hui ZHANG et Yiu-ming CHEUNG : The common self-polar triangle of concentric circles and its application to camera calibration. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4065–4072, 2015.
- [36] Haifei HUANG, Hui ZHANG et Yiu-ming CHEUNG : Homography estimation from the common self-polar triangle of separate ellipses. In IEEE Conference on Computer Vision Pattern Recognition (CVPR), pages 1737–1744, 2016.
- [37] Fredrik KAHL et Anders HEYDEN : Using conic correspondences in two images to estimate the epipolar geometry. In IEEE International Conference on Computer Vision (ICCV), 1998.
- [38] Masayuki KANBARA et Naokazu YOKOYA : Geometric and photometric registration for real-time augmented reality. In IEEE Proceedings of the 1st International Symposium on Mixed and Augmented Reality, page 279, 2002.
- [39] Juho KANNALA, Mikko SALO et Janne HEIKKILÄ : Algorithms for computing a planar homography from conics in correspondence. In British Machine Vision Conference (BMVC), pages 77–86, 2006.
- [40] Jun-Sik KIM, Pierre GURDJOS et In-So KWEON : Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 27(4), 2005.
- [41] Manfred KLOPSCHITZ et Dieter SCHMALSTIEG : Automatic reconstruction of wide-area fiducial marker models. In IEEE and ACM 6th International Symposium on Mixed and Augmented Reality (ISMAR), pages 71–74, 2007.
- [42] Laurent KNEIP, Davide SCARAMUZZA et Roland SIEGWART : A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2969–2976, 2011.
- [43] J. KOPP : Efficient numerical diagonalization of hermitian 3x3 matrices. International Journal of Modern Physics C, 2008.

- [44] Kevin KÖSER et Reinhard KOCH : Differential spatial resection-pose estimation using a single local image feature. In European Conference on Computer Vision (ECCV), pages 312–325. Springer, 2008.
- [45] Zuzana KUKELOVA, Martin BUJNAK et Tomas PAJDLA : Automatic generator of minimal problem solvers. In European Conference on Computer Vision (ECCV), pages 302–315. Springer, 2008.
- [46] Peter LIGHTBODY, Tomáš KRAJNÍK et Marc HANHEIDE : A versatile high-performance visual fiducial marker detection system with scalable identity encoding. In ACM Proceedings of the Symposium on Applied Computing, pages 276–282, 2017.
- [47] Haibin LING et Kazunori OKADA : Diffusion distance for histogram comparison. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 1, pages 246–253, 2006.
- [48] David G. LOWE : Distinctive image features from scale-invariant keypoints. International journal of computer vision (IJCV), 60(2):91–110, 2004.
- [49] David G LOWE et al. : Object recognition from local scale-invariant features. In IEEE International Conference on Computer Vision (ICCV), pages 1150–1157, 1999.
- [50] Matthew T. MASON : Mechanics of robotic manipulation. MIT press, 2001.
- [51] Fernando Eduardo Resende MATTIOLI et al. : Deep visual tracking for fiducial markers : a pilot study in telerehabilitation. 2019.
- [52] Kazuyuki MIYAZAWA, Koichi ITO, Takafumi AOKI, Koji KOBAYASHI et Hiroshi NAKAJIMA : An effective approach for iris recognition using phase-based image matching. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 30(10):1741–1756, 2008.
- [53] Jorge J MORÉ et Danny C SORENSEN : Computing a trust region step. SIAM Journal on Scientific and Statistical Computing, 4(3):553–572, 1983.
- [54] Pawan Kumar MUDIGONDA, CV JAWAHAR et PJ NARAYANAN : Geometric structure computation from conics. In Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), pages 9–14, 2004.
- [55] Rafael MUNOZ-SALINAS, Manuel J MARIN-JIMENEZ et Rafael MEDINA-CARNICER : Spm-slam : Simultaneous localization and mapping with squared planar markers. Pattern Recognition, 86:156–171, 2019.
- [56] Rafael MUNOZ-SALINAS, Manuel J. MARIN-JIMENEZ, Enrique YEGUAS-BOLIVAR et Rafael MEDINA-CARNICER : Mapping and localization from planar markers. Pattern Recognition, 73:158–171, 2018.
- [57] Raul MUR-ARTAL, Jose Maria Martinez MONTIEL et Juan D TARDOS : Orb-slam : a versatile and accurate monocular slam system. IEEE Transactions on Robotics, 31(5):1147–1163, 2015.
- [58] Leonid NAIMARK et Eric FOXLIN : Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. In IEEE Proceedings of the 1st International Symposium on Mixed and Augmented Reality (ISMAR), page 27, 2002.
- [59] Thanh Minh NGUYEN, Siddhant AHUJA et QM Jonathan WU : A real-time ellipse detection based on edge grouping. In IEEE International Conference on Systems, Man and Cybernetics, pages 3280–3286, 2009.

-
- [60] David NISTÉR : An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 26(6), 2004.
 - [61] David NISTER et Henrik STEWENIUS : Scalable recognition with a vocabulary tree. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 2, pages 2161–2168, 2006.
 - [62] Jiarong OU, Jin-Gang YU, Changxin GAO, Lichao XIAO et Zhifeng LIU : Robust and efficient ellipse fitting using tangent chord distance. In Asian Conference on Computer Vision (ACCV), pages 183–199. Springer, 2018.
 - [63] Alain PAGANI, Johannes KOEHLE et Didier STRICKER : Circular Markers for camera pose estimation. In 12th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), 2011.
 - [64] Théodore PAPADOPOULOU et Manolis IA LOURAKIS : Estimating the jacobian of the singular value decomposition : Theory and applications. In European Conference on Computer Vision (ECCV), pages 554–570. Springer, 2000.
 - [65] Viorica PĂTRĂUCEAN, Pierre GURDJOS et Rafael Grompone VON GIOI : A parameterless line segment and elliptical arc detector with enhanced ellipse fitting. In European Conference on Computer Vision (ECCV), pages 572–585. Springer, 2012.
 - [66] Mattis PAULIN, Matthijs DOUZE, Zaid HARCHAOUI, Julien MAIRAL, Florent PERONIN et Cordelia SCHMID : Local convolutional features with unsupervised training for image retrieval. In IEEE International Conference on Computer Vision (ICCV), pages 91–99, 2015.
 - [67] Ofir PELE et Michael WERMAN : Fast and robust earth mover’s distances. In IEEE International Conference on Computer Vision (ICCV), pages 460–467, 2009.
 - [68] Ofir PELE et Michael WERMAN : The quadratic-chi histogram distance family. In European conference on computer vision (ECCV), pages 749–762. Springer, 2010.
 - [69] Jun REKIMOTO : Matrix : A realtime object identification and registration method for augmented reality. In IEEE 3rd Asia Pacific Computer Human Interaction (Cat. No. 98EX110), pages 63–68, 1998.
 - [70] Paul L. ROSIN : Ellipse fitting using orthogonal hyperbolae and stirling’s oval. Graphical Models and Image Processing, 60(3):209–213, 1998.
 - [71] Junaed SATTAR, Eric BOURQUE, Philippe GIGUERE et Gregory DUDEK : Fourier tags : Smoothly degradable fiducial markers for use in human-robot interaction. In IEEE 4th Canadian Conference on Computer and Robot Vision (CRV), pages 165–174, 2007.
 - [72] Gerald SCHWEIGHOFER et Axel PINZ : Robust pose estimation from a planar target. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 28(12):2024–2030, 2006.
 - [73] David F SHANNO : Conditioning of quasi-newton methods for function minimization. Mathematics of computation, 24(111):647–656, 1970.
 - [74] Peter STURM : Algorithms for plane-based pose estimation. In IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), volume 1, 2000.
 - [75] Akihiro SUGIMOTO : A linear algorithm for computing the homography from conics in correspondence. Journal of Mathematical Imaging and Vision (JMIV), 13(2):115–130, 2000.
 - [76] Richard SZELISKI : Computer vision : algorithms and applications. Springer Science & Business Media, 2010.

- [77] Zygmunt L. SZPAK, Wojciech CHOJNACKI et Anton van den HENGEL : Guaranteed ellipse fitting with a confidence region and an uncertainty measure for centre, axes, and orientation. Journal of Mathematical Imaging and Vision (JMIV), 52(2), 2015.
- [78] Philip HS TORR et Andrew ZISSERMAN : Mlesac : A new robust estimator with application to estimating image geometry. Computer vision and image understanding (CVIU), 78(1), 2000.
- [79] Daniel WAGNER et Dieter SCHMALSTIEG : Artoolkitplus for pose tracking on mobile devices. In Computer Vision Winter Workshop, 2007.
- [80] John WANG et Edwin OLSON : Apriltag 2 : Efficient and robust fiducial detection. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 4193–4198, 2016.
- [81] Haiyuan WU, Qian CHEN et Toshikazu WADA : Conic-based algorithm for visual line estimation from one image. In IEEE Conference on Automatic Face and Gesture Recognition, 2004.
- [82] Yihong WU, Fulin TANG et Heping LI : Image-based camera localization : an overview. Visual Computing for Industry, Biomedicine, and Art, 1(1):1–13, 2018.
- [83] Zhengyou ZHANG : A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, 2000.
- [84] Yinqiang ZHENG, Wenjuan MA et Yuncai LIU : Another way of looking at monocular circle pose estimation. In IEEE International Conference on Image Processing (ICIP), pages 861–864, 2008.

Résumé

La localisation ou l'estimation de la pose d'une caméra à partir d'une vue est un problème essentiel dans de nombreuses applications, telles que le suivi d'une trajectoire de véhicule autonome ou la réalité augmentée.

La pose d'une caméra peut être estimée à partir de marqueurs construits spécifiquement pour cette tâche et qui peuvent être placés dans la scène. Ces marqueurs possèdent des caractéristiques qui permettent de les distinguer des autres objets de la scène. Ceci facilite leur détection dans une image et permet de les identifier plus facilement que des points d'intérêts. La correspondance entre la forme du marqueur et son image par une caméra contraint la pose de sorte qu'il est possible de retrouver les coordonnées du marqueur dans le repère de la caméra.

Les marqueurs circulaires sont particulièrement intéressants pour leur résilience à l'occultation partielle de l'image. En effet une ellipse peut être estimée à partir de seulement 5 points quelconques de son contour. Ainsi un marqueur partiellement occulté peut toujours être correctement localisé. Le problème est que l'image du contour circulaire n'est pas suffisante pour calculer la pose du marqueur, les marqueurs circulaires fonctionnent ainsi soit avec une caméra calibrée, soit en ajoutant d'autres primitives comme par exemple des cercles concentriques.

L'image du contour est pourtant souvent la seule information fiable dont on dispose dans certains cas d'utilisation. De plus, même avec une caméra calibrée, l'image d'un seul cercle n'est pas suffisante pour obtenir la pose de son plan de support. Il existe en effet une double ambiguïté qu'on ne peut distinguer sans information supplémentaire. En plus de cette ambiguïté, il existe une infinité de poses possibles pour le marqueur. Ces poses correspondent à l'ensemble des rotations autour de l'axe perpendiculaire au plan de support et passant par le centre du cercle. Ainsi nous pouvons distinguer quatre problématiques majeures traitées dans cette thèse :

- Comment prendre en compte des connaissances approximatives sur les paramètres de calibrage ?
- Quelles sont les conditions minimales de résolution de l'ambiguïté sur la pose du plan ?
- Comment calculer la rotation autour de l'axe du cercle ?
- Est-il possible d'identifier le marqueur en utilisant également son environnement ?

Notre première contribution répond à la première question. Pour cela nous proposons un modèle « par défaut » de la caméra. Ce modèle permet d'insérer nos incertitudes sur les paramètres intrinsèques de la caméra dans un algorithme qui recherche une solution photométrique au problème de la rectification métrique de l'image du marqueur. Les résultats démontrent que l'utilisation de ce modèle « par défaut » offre des perspectives encourageantes pour les marqueurs circulaires.

Notre seconde contribution porte sur l'ambiguïté de la pose du plan de support. Pour l'étudier nous proposons une nouvelle formulation du problème de l'estimation de la pose d'un cercle à partir de l'image de son contour. Cette formulation permet de faire le lien entre l'ambiguïté géométrique et les solutions algébriques obtenues. Cela nous amène à proposer une paramétrisation minimale de la pose qui fournit une condition simple de résolution de l'ambiguïté.

Enfin nous proposons une nouvelle méthode de localisation du marqueur circulaire à partir de points d'intérêts détectés dans son voisinage. La méthode repose sur une bibliothèque composée de vues de référence du marqueur prises hors ligne. Le marqueur est ensuite reconnu sur une nouvelle vue en utilisant conjointement l'image du cercle et des correspondances de points sur les vues de référence. Le modèle de pose est ensuite validé par une procédure de type RANSAC qui s'appuie sur une paramétrisation minimale par l'image d'un point 3D et d'un cercle. Par la suite nous pensons poursuivre les travaux sur la reconnaissance du marqueur notamment avec les nouvelles approches fondées sur l'apprentissage profond.

BIBLIOGRAPHIE
